PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(54) Title: METHODS AND COMPOSITIONS FOR THE TREATMENT OF BONE RESORPTION DISORDERS, INCLUDING OSTEOPOROSIS

(57) Abstract

The present invention relates to methods and compositions for the amelioration of symptoms caused by bone resorption disorders, including but not limited to osteoporosis, arthritides and periodontal disease, and damage caused by macrophage–mediated inflammatory processes. In one embodiment, the methods and compositions of the invention include methods and compositions for the specific inhibition of cathepsin K activity. In an additional embodiment, the methods and compositions of the invention include methods and compositions for the specific inhibition of cathepsin K activity coupled with specific inhibition of at least a second activity involved in the bone resorption and/or macrophage–mediated inflammatory processes. In a particular embodiment, the methods and compositions of the invention include methods and compositions for the specific inhibition of cathepsin K and cathepsin S activity.

METHODS AND COMPOSITIONS FOR THE TREATMENT
OF BONE RESORPTION DISORDERS, INCLUDING OSTEOPOROSIS

5

1.   INTRODUCTION

The present invention relates to methods and
compositions for the amelioration of symptoms caused by bone
10 resorption disorders, including but not limited to
osteoporosis, arthritides and periodontal disease, and damage
caused by macrophage-mediated inflammatory processes.

In one embodiment, the methods and compositions of the
invention include methods and compositions for the specific
15 inhibition of cathepsin K activity. In an additional
embodiment, the methods and compositions of the invention
include methods and compositions for the specific inhibition
of cathepsin K activity coupled with specific inhibition of
at least a second activity involved in the bone resorption
20 and/or macrophage-mediated inflammatory processes. In a
particular embodiment, the methods and compositions of the
invention include methods and compositions for the specific
inhibition of cathepsin K and cathepsin S activity.

25
2.   BACKGROUND

Bone remodeling is a dynamic process which does not
cease when longitudinal growth ceases, but continues
throughout an individual's lifetime. Remodeling is required
to preserve the mechanical strength of bone and involves both
30 bone resorption and deposition. In order to maintain a
constant adult skeletal mass, the rates of bone resorption
and formation must be equal. (For a review, see, e.g.,
Rasmussen, H. & Bordier, P., 1974, The Physiological and
Cellular Basis of Metabolic Bone Disease, Williams & Wilkins
35 Co., Baltimore).

Bone formation and resorption are mediated by the temporally and spatially related actions of specific bone cells. Among these bone cells are osteoblasts and osteoclasts (for a review, see Rodan, C.A., 1992, Bone

5 13:S3). Osteoblasts are involved in bone formation processes, including organic matrix synthesis and mineralization. Osteoclasts are polarized, multinucleate cells involved in bone resorption processes. Such processes include the creation of an acidic environment which leads to

10 bone demineralization, and the synthesis and production of proteinases which degrade the bone's organic matrix, 95% of which is Type I collagen (Krane, S.M. & Simon, L., 1994, in Scientific American Medicine, Rubenstein, E. & Felderman, D.D., eds., Scientific American, Inc., New York, pp. 1-26).

15. Among the proteinases that may be involved in bone organic matrix degradation and/or bone resorption are interstitial collagenase, and a complex group of acidic cysteine protease cathepsins of the papain superfamily (Kirschke, H. & Barrett, A.J., 1987, in Lysosomes: Their

20 Role in Protein Breakdown, Glaumann, H. & Ballard, F.J., eds., Academic Press, London, pp. 193-238; Kirschke, H. et al., 1986, Biochem. J. 264:467-473) including cathepsin B (Ferrara et al., 1990, FEBS Lett. 273:195-199; Chan et al., 1986, Proc. Natl. Acad. Sci. USA 83:7721-7725), S (Shi, G.-P.

25 et al., 1994, J. Biol. Chem. 269:11530-11536), L (Chauthan, S.S. et al., 1993, J. Biol. Chem. 268:1039-1045), O (Velasco, G. et al., 1994, J. Biol. Chem. 269:27136-27142) and a cathepsin which has been referred to alternatively as cathepsin O (unrelated to the above-mentioned O), OC2, O2, K

30 or X (Tezuka, K. et al., 1994, J. Biol. Chem. 269:1106-1109; Li, Y.P. et al., 1994, Mol. Biol. Cell 5:335a; Shi, G.-P. et al., 1995, FEBS Lett. 357:129-134; Inaoka, T. et al., 1995, Biochem. Biophys. Res. Commun. 206:89-96; Bromme, D. & Okamoto, K., 1995, Biol. Chem. Hoppe-Seyler 376:379-386), and

35 will be referred to herein as cathepsin K. Cathepsin K is expressed at high levels in osteoclast cells and in activated macrophages, but not in peripheral monocytes, despite the

fact that osteoclasts and monocytes are derived from a common progenitor. Further, a number of cathepsin B-like and cathepsin-L-like activities have been reported (Page et al., 1992, Biochim. Biophys. Acta 1116:57-66) from osteoclastomas.

5   The identification of which of the cathepsins may be involved in bone resorption has been problematic, given, for example, the that osteoclasts are rare cells and no acceptable osteoclast cell line has been established (Drake, F.H. et al., 1996, J. Biol. Chem. 271:12511-12516).

10 Nonetheless, there is evidence for cysteine proteases in bone resorption. For example, Delaisse et al. reported that cysteine protease inhibitors reduced bone resorption in a mouse bone organ culture system (Delaisse, J.-M. et al., 1980, Biochem. J. 192:365-368) and in calcium-deficient rats

15 (Delaisse, J.-M. et al., 1984, Biochem. Biophys. Res. Commun. 125:441-447), and classified the enzyme responsible as cathepsin B. In contrast, however, Drake et al. reports that cathepsins B, H and L are either expressed at very low levels or are absent in osteoclasts, and speculates that cathepsin

20 K, which, it is reported, is selectively and highly expressed in osteoclasts, is the important cathepsin in bone resorption (Drake, F.H. et al., 1996, J. Biol. Chem. 271:12511-12516). Thus, while one or more of the cathepsins may be involved in bone resorption and bone resorption disorders

25 such as those discussed below, conflicting data exist relating to the relative importance of the individual cathepsins.

Any interference with the normal remodeling process can produce skeletal disease, with the most common skeletal

30 disorders involving those associated with a decreased absolute skeletal mass, indicating that the rate of bone resorption exceeds the rate of bone formation. The most common of such disorders is osteoporosis. Other bone loss disorders include, for example, periodontal disease and

35 certain arthritides such as osteoarthritis.

Osteoporosis comprises several different bone resorption disorders associated with a decreased bone mass in the

absence of a mineralization defect, and is characterized by a
decreased skeletal mass, declining density of bone, loss of
trabeculations and a thinning of the cortices (see, e.g.,
Avioli, L.V. et al., 1990, in Metabolic Bone Disease and
5 Clinical Related Disorders, Avioli, L.V. & Krane, eds.,
Saunders Co., Philadelphia, p. 397; Melton, L.J. & Riggs,
B.L., 1989, in Clinical Disorders of Bone and Mineral
Metabolism, Kleerekoper, M. & Krane, S.M., eds., Mary Ann
Liebert, New York, p. 145; and Riggs, B.L. & Melton, L.J.
10 III, 1986, N. Engl. J. Med. 314:1676).
     The most common forms of osteoporosis, referred to
collectively as involutional osteoporosis, accompany aging
and occur particularly in women after menopause. In fact,
osteoporosis is the most significant underlying cause of
15 skeletal fractures in late middle age and elderly women.
Estrogen deficiency has been strongly implicated as a factor
in postmenopausal osteoporosis (Horowitz, M.C., 1993, Science
260:626).
     As opposed to osteoporosis, osteosclerotic disorders
20 involve an abnormal thickening or hardening of bone tissue.
Among such disorders is pycnodysostosis (PYCNO) a rare,
autosomal recessive trait characterized by osteosclerosis,
short stature, acro-osteolysis of distal phalanges, bone
fragility, clavicular dysplasia and skull deformities with
25 delayed suture closure (Maroteaux, P. & Lamy, M., 1962,
Presse Med. 70:999; Andren, L. et al., 1962, Acta. Chir.
Scand. 124:496). To date, no causative agent of the PYCNO
disorder has been identified.
     PYCNO is classified as an abnormality of primary
30 spongiosa resorption (Greenspan, A., 1991, Skeletal Radiol.
20:561). Histologically, osteoclast numbers, ruffled
borders, and clear zones are normal, but the region of
demineralized bone surrounding individual osteoclasts is
increased (Everts, V. et al., 1985, Calcif. Tissue Int.
35 37:25), and ultrastructural examination of osteoclasts reveal
occasional large abnormal cytoplasmic vacuoles containing
bone collagen fibrils. Such findings suggest that PYCNO

osteoclasts may be impaired in demineralizing bone and/or in degrading the organic matrix.

The gene responsible for PYCNO has not, to date, been identified. A genome-wide search for the PYCNO locus
5 involving a large, consanguineous Israeli Arab family with 16 affected relatives identified a genomic region that was homozygous-by-descent for all affected individuals (Gelb, B.D. et al., 1995, Nature Genet. 10:235). This region was localized to a pericentric portion of chromosome 1 with a lod
10 score of 11.72 at D1S498 and ancestral recombination events localized the PYCNO region to 4 cM from D1S442 to D1S305. Independent linkage analysis of a large Mexican family confirmed linkage to the chromosome 1 pericentric region and excluded the macrophage colony stimulating factor, CSF1,
15 which is defective in the op/op mouse model of osteoporosis (Polymeropoulos, M.H. et al., 1995, Nature Genet. 10:238).

No adequate treatments exist for bone disorders such as those described herein. With respect to osteoporosis, current treatment can, at best, attempt to prevent further
20 bone loss and, if possible, restore bone mass to some extent. Osteoporosis therapies can include estrogen (Albright, F. et al., 1941, JAMA 116:2465), calcium supplements (Prince, R.L. et al., 1991, N. Engl. J. Med. 325:1189) and vitamin D supplements (Chapuy, M.C. et al., 1992, N. Engl. J. Med.
25 327:1637). Further treatments can include sodium fluoride (Riggs, B.L. et al., 1992, N. Engl. J. Med. 327:620), androgen (Nagent de Deuxchaisnes, C., 1983, in Osteoporosis, a Multi-Disciplinary Problem, Royal Society of Medicine International Congress and Symposium Series No. 55, Academic
30 Press, London, p. 291)., calcitonin (Christiansen, C., 1992, Bone 13 (Suppl. 1):S35) or biphosphonate (Storm, T. et al., 1990, N. Engl. J. Med. 322:1265; Watts, N.B. et al;., 1990, N. Engl. J. Med. 323:73). Further, experimental therapies such as synthetic parathyroid hormone analogue (Hock, J.M. et
35 al., 1990, Endocrinology 127:1804) or thiazide diuretics (Cauley, J.A. et al., 1993, Ann. Intern. Med. 118:666) have also been attempted.

In summary, no agent responsible, *in vivo*, for bone
matrix degradation has, to date, been identified. A great
need exists for the definitive identification of targets for
the treatment of bone disorders, including bone resorption
5 disorders, such as osteoporosis and osteosclerotic disorders.

3. SUMMARY OF THE INVENTION

The present invention relates, first, to methods and
compositions for the amelioration of symptoms caused by bone
10 resorption disorders, including but not limited to
osteoporosis, arthritides such as osteoarthritis, and
periodontal disease, and damage caused by macrophage-mediated
inflammatory processes.

The present invention is based, in part, on the
15 Applicants' definitive demonstration that cathepsin K is
essential to bone resorption. This discovery represents the
first demonstration of an agent which acts *in vivo* to degrade
bone matrix. The present invention is further based on the
surprising discovery that the bone resorptive process may
20 require not only cathepsin K, but may also require the
presence of one or more additional proteases and/or other
components of the bone matrix hydrolyzing system.

In one embodiment, the methods of the invention comprise
methods for ameliorating symptoms of bone resorption
25 disorders and/or macrophage-mediated inflammatory damage via
methods for the specific inhibition of cathepsin K activity.

In an additional embodiment, the methods of the
invention include methods for the specific inhibition of at
least a second activity involved in bone resorption and/or
30 macrophage-mediated inflammatory damage processes. Such an
activity or activities can include, but is (are) not limited
to, activities that control cathepsin K expression,
activation, biosynthesis, processing, catalytic function,
stability, intracellular transport, osteoclast-specific
35 expression or enhanced expression, osteoclastic secretion,
macrophage expression or enhanced expression and/or
macrophage-mediated secretion. Such an activity or

activities can also include, for example, activities required
for coordinated matrix degradation, including, but not
limited to, activities involving another cathepsin, for
example, cathepsin S, L, B, D and/or O. Further, such
5 activities can include, but are not limited to, activities
involving availability or prior processing of a cathepsin
substrate for cathepsin hydrolysis; generation of active
osteoclasts from their monocyte/macrophage progenitors; and
function of osteoclasts in creating the acidic, sealed
10 microenvironment of the subosteoclastic space.

In a further embodiment, the methods of the invention
comprise methods for the specific inhibition of cathepsin K
activity coupled with specific inhibition of at least a
second activity, as described above, involved in bone
15 resorption and/or macrophage-mediated inflammatory damage
processes. For example, in a particular embodiment, the
methods of the invention include methods for the specific
inhibition of cathepsin K and cathepsin S activities.

It is also contemplated that polymorphisms and/or
20 mutations within the cathepsin K gene can affect bone density
in individuals carrying alleles containing such polymorphisms
and/or mutations. As such, it is a further object of the
invention to provide methods for the detection of bone-
density-affecting polymorphisms within the cathepsin K gene.

25 The compositions of the invention include, in one
embodiment, compositions, including pharmaceutical
compositions, for the specific inhibition of cathepsin K
activity. Such compositions of the present invention can
include, but are not limited to, inhibitors of cathepsin K
30 gene activity, such as, for example, cathepsin K antisense,
triple helix and/or ribozyme molecules, and inhibitors of
cathepsin K enzymatic activity.

Cathepsin K specific inhibitors can also include, but
are not limited to, specific inhibitors belonging to the
35 following, preferably peptidyl, classes of compounds:
fluoromethyl ketones, vinyl sulfones; peptide aldehydes,
nitriles, α-ketocarbonyl compounds, including, for example,

α-diketones, α-keto esters, α-ketoamides, and α-ketoacids, halomethyl ketones, diazomethyl ketones, (acyloxy)-methyl ketones, ketomethylsulfonium salts and epoxysuccinyl compounds.

5   In a further embodiment, the compositions of the invention include compositions, including pharmaceutical compositions, for the specific inhibition of cathepsin K activity and at least a second activity, as described above, involved in bone resorption and/or macrophage-mediated

10 inflammatory damage processes. For example, in a particular embodiment, the compositions include compositions, including pharmaceutical compositions, for the specific inhibition of cathepsin K and cathepsin S activities.

        In instances in which the additional activity or

15 activities is (are) additional cathepsin activities, the additional cathepsin specific inhibitors can also include, but are not limited to, specific inhibitors belonging to the following, preferably peptidyl, classes of compounds: fluoromethyl ketones, vinyl sulfones, peptide aldehydes,

20 nitriles, α-ketocarbonyl compounds, including, for example, α-diketones, α-keto esters, α-ketoamides, and α-ketoacids, halomethyl ketones, diazomethyl ketones, (acyloxy)-methyl ketones, ketomethylsulfonium salts and epoxysuccinyl compounds.

25   The present invention additionally relates to methods and compositions for the amelioration of symptoms caused by osteosclerotic bone disorders, including but not limited to pycnodysostosis (PYCNO).

        Further, the present invention relates to isolated

30 osteoclast cells and cell lines deficient for cathepsin K and/or other cathepsin activities. Still further, the present invention relates to genetically engineered nonhuman animals deficient for cathepsin K and other cathepsin activities.

35   The Example presented in Section 6 demonstrates that the osteosclerotic bone disorder PYCNO is caused by a deficiency of cathepsin K activity. By elucidating the cause of the

natural PYCNO disorder, the present invention definitively
identifies for the first time, an agent, cathepsin K, which
degrades bone matrix *in vivo* and which is essential to bone
resorption.

5   The Example presented in Section 7, below, indicates
that a mutation or polymorphism within at least one other
(i.e., second) gene locus appears to predispose individuals
to PYCNO, indicating that the disease requires the presence
of at least two mutations, one present in the cathepsin K

10   gene and at least one additional mutation present in a second
gene involved in the bone resorption process. The example
presented in Section 8 describes the genomic structure of the
human cathepsin K gene.

         Example 9 presents the isolation and characterization of

15   the murine cathepsin K gene. The isolation of the murine
cathepsin K gene will facilitate the production of
genetically engineered transgenic mice deficient for
cathepsin K. Such mice will serve as an important animal
model system for studying cathepsin K gene expression and

20   function and development of methods and compositions for
amelioration of symptoms caused by bone resorption disorders.

         4.   BRIEF DESCRIPTION OF THE FIGURES
         FIG. 1.   Pycnodysostosis (PYCNO) critical region at

25   chromosome 1q21. The microsatellite markers flanking and
within the PYCNO critical region are shown in the order
established previously (Gelb, B. D. et al., 1996, Hum. Genet.
98:141-144). Two overlapping CEPH megaYAC clones which were
positive for the cathepsin S and K STSs are indicated and

30   anchored by *D1S498*.
         FIG. 2.   Cathepsin K cDNA and polypeptide with PYCNO
mutations.   The top line represents the cathepsin K exon
structure; the second line represents the cDNA with
initiation (ATG) and stop (TGA) codons as indicated.   The

35   locations of the eight mutations are indicated below the
cDNA, showing both the affected codon and the predicted amino
acid alteration.   The normal cathepsin K polypeptide is shown

with the pre-, pro-, and mature peptides. The three residues
(C, H, N) in the active site conserved among all cysteine
protease cathepsins in the papain superfamily are indicated.
The K52X and R241X mutations are nonsense mutations

5 resulting in truncated proteins, the E79G, G146R, Y212C,
A277E and R312G mutations are missense mutations and the
X330W mutation is shown with the predicted elongation of the
mature peptide by 19 residues. At the right are listed the
families in which the respective mutations were found.

10 (Port. = Portugal; MSP = Mount Sinai PYCNO patient; IP =
Israeli PYCNO patient).

        FIG. 3A-3B. Transient expression of the X330W mutation
in cathepsin K. (A) Wild type and X330W cDNAs were subcloned
into the pcDNA-1 expression vector (Invitrogen) (Shi, G.-P.

15 et al., 1995, FEBS Lett. 357:129; Bromme, D. & Okamoto, K.,
1995, Biol. Chem. Hoppe-Seyler 376:37; Tezuka, K., et al.,
1994, J. Biol. Chem. 269:1106; Inaoka, T., et al., 1995,
Biochem. BioPhys. Res. Commun. 206:89) and confirmed by PCR,
restriction enzyme digestions and dideoxy chain termination

20 sequencing. Both constructs were transfected into 293 cells
using LipofectAMINE (GibcoBRL) along with 3 μg pAdVAmtage™
(Promega) which contains the adenovirus virus-associated I
RNA (VAI) to enhance translation (Kaufman, R.J. & Murtha, P.,
1987, Mol. Cell. Biol. 7:1568). Cells were harvested after

25 48 h post-transfection and assayed by immunoblotting. Blots
were developed with monospecific polyclonal anti-human
cathepsin K antibodies raised by injection of cathepsin K-
maltose binding protein fusion protein into rabbits and
purified by elution from immobilized antigen as described

30 (Munger, J. S. et al., 1995, Biochem. J., 311:299). Lane 1
shows results with nontransfected cells. Lanes 2-3, 4-5, and
6-7 show results with cells transfected with 3, 10 or 15 μg,
respectively, of wild-type (W) or mutant (M) cDNA. (B) Total
RNA was isolated from both transfected and nontransfected 293

35 cells and 40 μg of total RNA was separated on a 2.4% agarose
gel. Northern blot was hybridized with [^{32}P]-labeled probes
of both cathepsin K and VAI. Transcripts of both cathepsin K

and VAI are indicated. Lanes 1 to 7 have the same order as
in (A).

  FIG. 4. Genomic organization of the human cathepsin K
gene. The eight exons are shown as black boxes with the
5 translation initiation ATG and TGA stop codons indicated.
The seven introns are indicated by the thick black lines.
The cumulative size of the cathepsin K genomic sequence is
shown on the scale below.

  FIG. 5. Identification of the human cathepsin K
10 transcription initiation sites by primer extension. Fifty
micrograms of human brain total RNA was used for the primer
extension (PE) reaction with an antisense primer from
nucleotides +84 to +60. The primer extension products were
determined to be 124, 186, and 253 bp in length by denaturing
15 polyacrylamide gel electrophoresis using the adjacent DNA
sequencing reaction (GATC) as a size marker. this
corresponded to expected 5' UTRs of 40, 102, and 169 nt,
respectively.

  FIG. 6. Sequence of human cathepsin K exon 1 and
20 upstream flanking region with comparison to the published 5'
untranslated regions from cathepsin K cDNA clones. Exon 1
(capitalized) and upstream flanking (lowercase) sequence was
derived from PAC 74e16. The three transcription initiation
sites from primer extension analysis are indicated in bold
25 type. The 5' untranslated regions from four different
cathepsin k cDNA clones are aligned to exon 1 with identical
bases indicated. The sense primer sequence, CTSK43, used for
PCR amplification from first strand cDNA (see Section 8,
below) is underlined. Two putative AP1 binding sites in the
30 upstream flanking region are also underlined.

  FIG. 7. Nucleotide and predicted amino acid sequences
of the mouse cathepsin K cDNA. Nucleotide and amino acid
numbering are shown to the right and left, respectively. The
A nucleotide of the initiation ATG is designated nt 136 and
35 the N-terminal Met of the signal peptide is amino acid 1.
The first amino acids of the proregion and the mature enzyme
following the putative cleavage sites are indicated as Pro

and Mature, respectively. The two potential *N*-glycosylation
consensus sequences are underlined in the coding region. The
active site amino acids (C139, H276, and N296) are indicated
by boldface type. Two AATAAA polyadenylation signals and one
5 consensus recognition sequence (CACTG) for the U4 small
nuclear ribonucleoprotein are underlined in the 3'
untranslated region.

FIG. 8.   Amino acid alignment of mouse cathepsin K with
the human, rabbit, and chicken homologues. The numbering is
10 based on the mouse sequence with the initial methionine
assigned as 1. The three active site amino acids (C139,
H276, and N296) are indicated by asterisks. Sequence
identity is indicated by the shaded boxes.

FIG. 9.   Genomic organization of the mouse cathepsin K
15 gene. The eight exons are shown as block boxes with the
translation initiation ATG and TGA stop codons indicated.
The seven introns are indicated by the thick black lines.
The cumulative size of the cathepsin K genomic sequence is
shown on the scale below. The size of intron 1 was not
20 determined, as indicated by the hatched lines. The portions
of the coding region corresponding to the presignal,
propeptide, and mature enzyme are indicated.

5.   DETAILED DESCRIPTION OF THE INVENTION

25   5.1. COMPOUNDS FOR THE AMELIORATION OF SYMPTOMS
             ASSOCIATED WITH BONE RESORPTION DISORDERS

Described herein are compositions, including
pharmaceutical compositions, which can be utilized for the
amelioration of symptoms associated with bone resorption
30 disorders, including, but not limited to, osteoporosis,
arthritides such as osteoarthritis, and periodontal disease,
and damage caused by macrophage-mediated inflammatory
processes. Such compositions can include, but are not
limited to, small peptides, small organic molecules,
35 antisense, ribozyme and triple helix molecules.

Among the compositions of the invention are
compositions, including pharmaceutical compositions, which

represent specific inhibitors of cathepsin K activity.
Pharmaceutical compositions of the invention are described in
Section 5.2.1, below.

The term "specific", as used herein, refers to a
5 composition that reduces the level of cathepsin K enzymatic
activity to a greater extent than it reduces the enzymatic
activity levels of other cysteine proteases, serine
proteases, metalloproteases and aspartyl proteases.
Preferably, an inhibitor of this class reduces the level of
10 cathepsin K activity while not appreciably affecting the
enzymatic activity of other such proteases. While not
wishing to be bound by a particular mechanism, the inhibitors
can act, for example, by lowering the level of cathepsin K
gene expression or may act directly on the cathepsin K enzyme
15 to inhibit or suppress enzymatic activity.

In instances wherein the inhibitor of this class is to
be used in conjunction with the inhibition of at least a
second activity involved in bone resorption and/or
macrophage-mediated inflammatory damage processes, such an
20 inhibitor is still considered a "specific" inhibitor when it
reduces the level of an additional activity or activities to
a greater extent than it reduces other activities, such as,
for example, other cysteine protease activities. For
example, in a particular embodiment wherein cathepsin K and
25 cathepsin S activities are to be inhibited, a specific
inhibitor of this class can refer to a composition which
inhibits both cathepsin K and cathepsin S activities to a
greater extent than it reduces the activities of other
cysteine proteases.
30   Among the cathepsin K specific inhibitors are those
molecules identified via the methods described, below, in
Section 5.4. Additionally such inhibitors can include, for
example, antisense, ribozyme and triple helix molecules, as
described in Section 5.1.1, below.
35   Cathepsin K specific inhibitors can also include, but
are not limited to, specific inhibitors belonging to the
following, preferably peptidyl, classes of compounds:

- 13 -

fluoromethyl ketones, vinyl sulfones, peptide aldehydes,
nitriles, α-ketocarbonyl compounds, including, for example,
α-diketones, α-keto esters, α-ketoamides, and α-ketoacids,
halomethyl ketones, diazomethyl ketones, (acyloxy)-methyl
5 ketones, ketomethylsulfonium salts and epoxysuccinyl
compounds.

Cathepsin K specific inhibitors can further include, but
are not limited to, cathepsin K proenzyme peptides. Such pro
enzyme peptides comprise the "propart" region of cathepsin K
10 protein independent of the mature cathepsin K protein
sequence. The cathepsin K propart generally refers to the 98
amino acid sequence at the N-terminus of the cathepsin K
proenzyme which remains after cleavage of the "pre" leader
sequence. (See, e.g., Shi, G.-P. et al., 1995, FEBS Lett.
15 357:129-134, which is incorporated herein by reference in its
entirety.)

The compositions of the invention can further include
specific inhibitors of at least a second activity involved in
bone resorption and macrophage-mediated inflammatory damage
20 processes. The compounds in this class of inhibitors
include, but are not limited to, specific inhibitors of a
gene product involved in cathepsin K expression,
biosynthesis, processing, activation, catalytic function,
intracellular transport, stability, osteoclast-specific
25 expression or enhanced expression, osteoclastic secretion,
macrophage expression or enhanced expression and/or
macrophage secretion.

Further, the compounds in this class include, but are
not limited to, specific inhibitors of a gene product
30 involved in the availability or prior processing of the
cathepsin K substrate(s) for hydrolysis by cathepsin K.

Still further, the compounds in this class include, but
are not limited to, inhibitors of: another cathepsin, such
as, for example, cathepsin S, L, B, D and/or O; a gene
35 product involved in the expression, enzymatic activity and/or
substrate availability of such an additional cathepsin; the
generation of active osteoclasts from their

monocyte/macrophage progenitors; and/or a gene product
involved in the function of osteoclasts such as, for example,
in creating the acidic, sealed microenvironment of the
subosteoclastic space.

5   The term "specific", as used herein, refers to a
composition that reduces (that is, inhibits or suppresses)
the activity level or levels of interest, to a greater extent
than it reduces the remainder of the activities listed above.
For example, if the composition is an inhibitor of a second
10 activity relating to cathepsin K activity, then the inhibitor
is a specific one if its effect is to reduce the level of
cathepsin K activity to a greater extent than it reduces the
activity of other cysteine proteases, including serine
proteases, aspartyl proteases, etc.

15   When an inhibitor of this class is to be utilized
coupled with an inhibitor of cathepsin K activity, it is to
still be considered a specific inhibitor if it inhibits both
the second activity of interest (cathepsin S activity, for
example) as well as inhibiting cathepsin K activity to a
20 greater extent than it reduces the activity of the remainder
of the activities listed above. Further, an inhibitor of
this class is still considered a specific inhibitor if the
composition inhibits both cathepsin K activity and the second
activity of interest (such as, for example, cathepsin S
25 activity).

      5.1.1. Antisense, Ribozyme and Triple Helix Inhibitors
      Among the inhibitors of the invention are nucleic acid
antisense, ribozyme and/or triple helix molecules which act
30 to inhibit expression of genes involved in one or more of the
activities relating to bone resorption and/or macrophage-
mediated inflammatory damage processes. Such inhibitors can
be utilized in methods for the amelioration of symptoms
caused by bone resorption disorders and/or macrophage-
35 mediated inflammatory damage. For purposes of clarity,
antisense, triple helix and ribozyme inhibitors of cathepsin

K gene expression will frequently be used here as an example
and not by way of limitation.

Antisense or ribozyme approaches can be utilized to
inhibit or prevent translation of mRNA transcripts; triple
5 helix approaches to inhibit transcription of the gene of
interest itself.

Antisense approaches involve the design of
oligonucleotides (either DNA or RNA) that are complementary
to mRNA, cathepsin K mRNA, for example. The antisense
10 oligonucleotides bind to the complementary mRNA transcripts
and prevent translation. Absolute complementarity, although
preferred, is not required. A sequence "complementary" to a
portion of an RNA, as referred to herein, means a sequence
having sufficient complementarity to be able to hybridize
15 with the RNA, forming a stable duplex; in the case of double-
stranded antisense nucleic acids, a single strand of the
duplex DNA may thus be tested, or triplex formation may be
assayed. The ability to hybridize will depend on both the
degree of complementarity and the length of the antisense
20 nucleic acid. Generally, the longer the hybridizing nucleic
acid, the more base mismatches with an RNA it may contain and
still form a stable duplex (or triplex, as the case may be).
One skilled in the art can ascertain a tolerable degree of
mismatch by use of standard procedures to determine the
25 melting point of the hybridized complex.

Oligonucleotides that are complementary to the 5' end of
the message, e.g., the 5' untranslated sequence up to and
including the AUG initiation codon, should work most
efficiently at inhibiting translation. However, sequences
30 complementary to the 3' untranslated sequences of mRNAs have
recently been shown to be effective at inhibiting translation
of mRNAs as well. See generally, Wagner, R., 1994, Nature
372:333-335. Thus, taking cathepsin K as an example,
oligonucleotides complementary to the 5'-non-translated, non-
35 coding region, as shown in FIG. 6, and/or the 3'-non
translated, non-coding region (see, e.g., Tezuka, K. et al.,
1994, J. Biol. Chem. 269:1106-1109, which is incorporated

- 16 -

herein by reference in its entirety) the cathepsin K gene
could be used in an antisense approach to inhibit translation
of endogenous cathepsin K mRNA. Oligonucleotides
complementary to the 5' untranslated region of the mRNA

5 should include the complement of the AUG start codon.
Antisense oligonucleotides complementary to mRNA coding
regions are less efficient inhibitors of translation but
could be used in accordance with the invention. Whether
designed to hybridize to the 5'-, 3'- or coding region of

10 cathepsin K mRNA, antisense nucleic acids should be at least
six nucleotides in length, and are preferably
oligonucleotides ranging from 6 to about 50 nucleotides in
length. In specific aspects the oligonucleotide is at least
10 nucleotides, at least 17 nucleotides, at least 25

15 nucleotides or at least 50 nucleotides.

        Regardless of the choice of target sequence, it is
preferred that *in vitro* studies are first performed to
quantitate the ability of the antisense oligonucleotide to
inhibit gene expression. It is preferred that these studies

20 utilize controls that distinguish between antisense gene
inhibition and nonspecific biological effects of
oligonucleotides. It is also preferred that these studies
compare levels of the target RNA or protein with that of an
internal control RNA or protein. Additionally, it is

25 envisioned that results obtained using the antisense
oligonucleotide are compared with those obtained using a
control oligonucleotide. It is preferred that the control
oligonucleotide is of approximately the same length as the
test oligonucleotide and that the nucleotide sequence of the

30 oligonucleotide differs from the antisense sequence no more
than is necessary to prevent specific hybridization to the
target sequence.

        The oligonucleotides can be DNA or RNA or chimeric
mixtures or derivatives or modified versions thereof, single-

35 stranded or double-stranded. The oligonucleotide can be
modified at the base moiety, sugar moiety, or phosphate
backbone, for example, to improve stability of the molecule,

hybridization, etc. The oligonucleotide may include other
appended groups such as peptides (e.g., for targeting host
cell receptors in vivo), or agents facilitating transport
across the cell membrane (see, e.g., Letsinger et al., 1989,
5 Proc. Natl. Acad. Sci. U.S.A. 86:6553-6556; Lemaitre et al.,
1987, Proc. Natl. Acad. Sci. 84:648-652; PCT Publication No.
WO88/09810, published December 15, 1988) or the blood-brain
barrier (see, e.g., PCT Publication No. WO89/10134, published
April 25, 1988), hybridization-triggered cleavage agents.
10 (See, e.g., Krol et al., 1988, BioTechniques 6:958-976) or
intercalating agents. (See, e.g., Zon, 1988, Pharm. Res.
5:539-549). To this end, the oligonucleotide may be
conjugated to another molecule, e.g., a peptide,
hybridization triggered cross-linking agent, transport agent,
15 hybridization-triggered cleavage agent, etc.

       The antisense oligonucleotide may comprise at least one
modified base moiety which is selected from the group
including but not limited to 5-fluorouracil, 5-bromouracil,
5-chlorouracil, 5-iodouracil, hypoxanthine, xantine,
20 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil,
5-carboxymethylaminomethyl-2-thiouridine,
5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-
galactosylqueosine, inosine, N6-isopentenyladenine,
1-methylguanine, 1-methylinosine, 2,2-dimethylguanine,
25 2-methyladenine, 2-methylguanine, 3-methylcytosine,
5-methylcytosine, N6-adenine, 7-methylguanine,
5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil,
beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil,
5-methoxyuracil, 2-methylthio-N6-isopentenyladenine,
30 uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil,
queosine, 2-thiocytosine, 5-methyl-2-thiouracil,
2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-
5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v),
5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl)
35 uracil, (acp3)w, and 2,6-diaminopurine.

       The antisense oligonucleotide may also comprise at least
one modified sugar moiety selected from the group including

but not limited to arabinose, 2-fluoroarabinose, xylulose,
and hexose.

In yet another embodiment, the antisense oligonucleotide
comprises at least one modified phosphate backbone selected
5 from the group consisting of a phosphorothioate, a
phosphorodithioate, a phosphoramidothioate, a
phosphoramidate, a phosphordiamidate, a methylphosphonate, an
alkyl phosphotriester, and a formacetal or analog thereof.

In yet another embodiment, the antisense oligonucleotide
10 is an α-anomeric oligonucleotide. An α-anomeric
oligonucleotide forms specific double-stranded hybrids with
complementary RNA in which, contrary to the usual β-units,
the strands run parallel to each other (Gautier et al., 1987,
Nucl. Acids Res. 15:6625-6641). The oligonucleotide is a 2'-
15 0-methylribonucleotide (Inoue et al., 1987, Nucl. Acids Res.
15:6131-6148), or a chimeric RNA-DNA analogue (Inoue et al.,
1987, FEBS Lett. 215:327-330).

Oligonucleotides of the invention may be synthesized by
standard methods known in the art, e.g. by use of an
20 automated DNA synthesizer (such as are commercially available
from Biosearch, Applied Biosystems, etc.). As examples,
phosphorothioate oligonucleotides may be synthesized by the
method of Stein et al. (1988, Nucl. Acids Res. 16:3209),
methylphosphonate oligonucleotides can be prepared by use of
25 controlled pore glass polymer supports (Sarin et al., 1988,
Proc. Natl. Acad. Sci. U.S.A. 85:7448-7451), etc.

While antisense nucleotides complementary to the
cathepsin K coding region sequence could be used, those
complementary to the transcribed untranslated region are most
30 preferred.

The antisense molecules should be delivered to cells
which express cathepsin K in vivo, e.g., osteoclast and/or
macrophage cells. A number of methods have been developed
for delivering antisense DNA or RNA to cells; e.g., antisense
35 molecules can be injected directly into the tissue site, or
modified antisense molecules, designed to target the desired
cells (e.g., antisense linked to peptides or antibodies that

- 19 -

specifically bind receptors or antigens expressed on the
target cell surface) can be administered systemically.

However, it is often difficult to achieve intracellular
concentrations of the antisense sufficient to suppress
5 translation of endogenous mRNAs. Therefore a preferred
approach utilizes a recombinant DNA construct in which the
antisense oligonucleotide is placed under the control of a
strong pol III or pol II promoter. The use of such a
construct to transect target cells in the patient will result
10 in the transcription of sufficient amounts of single stranded
RNAs that will form complementary base pairs with the
endogenous cathepsin K transcripts and thereby prevent
translation of the cathepsin K mRNA. For example, a vector
can be introduced in vivo such that it is taken up by a cell
15 and directs the transcription of an antisense RNA. Such a
vector can remain episomal or become chromosomally
integrated, as long as it can be transcribed to produce the
desired antisense RNA. Such vectors can be constructed by
recombinant DNA technology methods standard in the art.
20 Vectors can be plasmid, viral, or others known in the art,
used for replication and expression in mammalian cells.
Expression of the sequence encoding the antisense RNA can be
by any promoter known in the art to act in mammalian,
preferably human cells. Such promoters can be inducible or
25 constitutive. Such promoters include but are not limited to:
the SV40 early promoter region (Bernoist and Chambon, 1981,
Nature 290:304-310), the promoter contained in the 3' long
terminal repeat of Rous sarcoma virus (Yamamoto et al., 1980,
Cell 22:787-797), the herpes thymidine kinase promoter
30 (Wagner et al., 1981, Proc. Natl. Acad. Sci. U.S.A. 78:1441-
1445), the regulatory sequences of the metallothionein gene
(Brinster et al., 1982, Nature 296:39-42), etc. Any type of
plasmid, cosmid, YAC or viral vector can be used to prepare
the recombinant DNA construct which can be introduced
35 directly into the tissue site; e.g., the choroid plexus or
hypothalamus. Alternatively, viral vectors can be used which
selectively infect the desired tissue; (e.g., for brain,

- 20 -

herpesvirus vectors may be used), in which case administration may be accomplished by another route (<u>e.g.</u>, systemically).

Ribozyme molecules designed to catalytically cleave
5 cathepsin K mRNA transcripts can also be used to prevent translation of cathepsin K mRNA and expression of cathepsin K enzyme. (See, <u>e.g.</u>, PCT International Publication .WO90/11364, published October 4, 1990; Sarver et al., 1990, Science <u>247</u>:1222-1225). While ribozymes that cleave mRNA at
10 site specific recognition sequences can be used to destroy cathepsin K mRNAs, the use of hammerhead ribozymes is preferred. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The sole requirement is that the
15 target mRNA have the following sequence of two bases: 5'-UG-3'. The construction and production of hammerhead ribozymes is well known in the art and is described more fully in Haseloff and Gerlach, 1988, Nature, <u>334</u>:585-591. There are many potential hammerhead ribozyme cleavage sites within the
20 nucleotide sequence of human cathepsin K, the sequence of which is well known to those of skill in the art. Preferably the ribozyme is engineered so that the cleavage recognition site is located near the 5' end of the cathepsin K mRNA; <u>i.e.</u>, to increase efficiency and minimize the intracellular
25 accumulation of non-functional mRNA transcripts.

The ribozymes of the present invention also include RNA endoribonucleases (hereinafter "Cech-type ribozymes") such as the one which occurs naturally in Tetrahymena thermophila (known as the IVS, or L-19 IVS RNA) and which has been
30 extensively described by Thomas Cech and collaborators (Zaug, et al., 1984, Science, <u>224</u>:574-578; Zaug and Cech, 1986, Science, <u>231</u>:470-475; Zaug, et al., 1986, Nature, <u>324</u>:429-433; published International patent application No. WO 88/04300 by University Patents Inc.; Been and Cech, 1986,
35 Cell, <u>47</u>:207-216). The Cech-type ribozymes have an eight base pair active site which hybridizes to a target RNA sequence whereafter cleavage of the target RNA takes place.

The invention encompasses those Cech-type ribozymes which
target eight base-pair active site sequences that are present
in cathepsin K.

As in the antisense approach, the ribozymes can be
5 composed of modified oligonucleotides (e.g. for improved
stability, targeting, etc.) and should be delivered to cells
which express the cathepsin K gene *in vivo*, e.g., osteoclasts
and/or macrophage cells. A preferred method of delivery
involves using a DNA construct "encoding" the ribozyme under
10 the control of a strong constitutive pol III or pol II
promoter, so that transfected cells will produce sufficient
quantities of the ribozyme to destroy endogenous cathepsin K
messages and inhibit translation. Because ribozymes, unlike
antisense molecules, are catalytic, a lower intracellular
15 concentration is required for efficiency.

Alternatively, endogenous cathepsin K gene expression
can be reduced by targeting deoxyribonucleotide sequences
complementary to the regulatory region of the cathepsin K
gene (*i.e.*, the cathepsin K promoter and/or enhancers) to
20 form triple helical structures that prevent transcription of
the cathepsin K gene in target cells in the body. (See
generally, Helene, C. 1991, Anticancer Drug Des., 6(6):569-
84; Helene, C., et al., 1992, Ann, N.Y. Acad. Sci., 660:27-
36; and Maher, L.J., 1992, Bioassays 14(12):807-15). Such
25 cathepsin K regulatory sequences can readily be obtained by
one of ordinary skill in the art by utilizing standard
molecular biological techniques coupled with the cathepsin K
gene 5' untranslated nucleotide sequence depicted herein in
FIG. 6.
30

5.2. METHODS FOR AMELIORATION OF SYMPTOMS ASSOCIATED WITH BONE RESORPTION DISORDERS

This Section describes methods for the amelioration of
bone resorption disorders, including, but not limited to
35 osteoporosis, arthritides such as osteoarthritis, and
periodontal disease, and damage caused by macrophage-mediated
inflammatory processes. Macrophage-mediated inflammatory

damage can include, but it is not limited to, macrophage-
mediated periodontal disease damage and osteoarthritic
damage. Such macrophage-mediated inflammatory damage can
further include damage in addition to specifically bone-
5 related damage, such as, for example, lung injury associated
with emphysema, and injury associated with an accumulation of
macrophages at the site of atherosclerotic plaques. In
general, the methods of the invention can be utilized for
ameliorating abnormal or excessive damage which occurs due to
10 macrophage accumulation at sites of injury.

    The methods of the invention further include methods for
the specific inhibition of at least a second activity
involved in bone resorption and/or macrophage-mediated
inflammatory damage processes. Such an activity or
15 activities can include, but is (are) not limited to,
activities that control cathepsin K expression, activation,
biosynthesis, processing, catalytic function, stability,
intracellular transport, osteoclast-specific expression or
enhanced expression, osteoclastic secretion, macrophage
20 expression or enhanced expression and/or macrophage-mediated
secretion. Such an activity or activities can also include,
for example, activities required for coordinated matrix
degradation, including, but not limited to, activities
involving another cathepsin, for example, cathepsin S, L, B,
25 D and/or O. Further, such activities can include, but are
not limited to, activities involving availability or prior
processing of a cathepsin substrate for cathepsin hydrolysis;
generation of active osteoclasts from their
monocyte/macrophage progenitors; and function of osteoclasts
30 in creating the acidic, sealed microenvironment of the
subosteoclastic space.

    In a further embodiment, the methods of the invention
include methods for the specific inhibition of cathepsin K
activity coupled with specific inhibition of at least a
35 second activity, as described above, involved in bone
resorption and/or macrophage-mediated inflammatory damage
processes. For example, in a particular embodiment, the

methods of the invention include methods for the specific
inhibition of cathepsin K and cathepsin S activities.

        A method for ameliorating symptoms of a bone resorption
disorder and/or symptoms of macrophage-mediated inflammatory
5   damage, can, for example, comprise: contacting a cell with a
cathepsin K inhibitor for a time sufficient to inhibit
cathepsin K activity so that symptoms of the disorder or
damage are ameliorated. The cathepsin K inhibitor can
include, but is not limited to, the cathepsin K inhibitor
10 compositions, including pharmaceutical compositions,
described in Section 5.1, above.

        A method for ameliorating symptoms of a bone resorption
disorder and/or symptoms cf macrophage-mediated inflammatory
damage, can also, for example, comprise: contacting a cell
15 with an inhibitor of at least a second activity involved in a
bone resorption process and/or a macrophage-mediated
inflammatory damage process, for a time sufficient to inhibit
the activity so that symptoms of the bone resorption disorder
are ameliorated. The inhibitor can include, but is not
20 limited to the inhibitor compositions, including
pharmaceutical compositions, described in Section 5.1, above.

        A method for ameliorating symptoms of a bone resorption
disorder and/or symptoms of macrophage-mediated inflammatory
damage, can, for example, still further, comprise:
25 contacting a cell with a cathepsin K inhibitor and an
inhibitor of at least a second activity involved in a bone
resorption process and/or a macrophage-mediated inflammatory
damage process, for a time sufficient to inhibit the activity
so that symptoms of the bone resorption disorder are
30 ameliorated. In a particular embodiment of such a method,
the second activity is a cathepsin S activity.

        It is envisioned that, in certain instances, a single
composition or pharmaceutical composition will represent both
the cathepsin K inhibitor and the inhibitor of the at least
35 second activity. In a particular embodiment, such an
inhibitor inhibits both cathepsin K and cathepsin S
activities.

It should be noted that in addition to such a second
activity, the methods of the invention also include
administration of inhibitors of additional (i.e., third,
fourth, etc.) activities in combination for the amelioration
5 of said symptoms.

When a combination of inhibitors is being utilized as
part of such methods of the invention, it is to be understood
that the inhibitors can be administered simultaneously or
sequentially. If administered simultaneously, the inhibitors
10 can be administered in any order most beneficial to
ameliorating symptoms of the disorder or damage. One of
reasonable skill in the art will, depending on the particular
inhibitors and disorder or damage being treated, readily
appreciate the most advantageous administration strategy to
15 follow.

5.2.1. PHARMACEUTICAL PREPARATIONS AND METHODS OF ADMINISTRATION

Compositions to be administered as part of methods for
20 ameliorating symptoms of bone resorption disorders and/or
macrophage-mediated inflammatory damage are to be
administered to a patient at therapeutically effective doses
to treat or ameliorate symptoms of such disorders. A
therapeutically effective dose refers to that amount of the
25 compound sufficient to result in amelioration of symptoms of
bone resorption disorders and/or macrophage-mediated
inflammatory damage.

The inhibitory compositions of the invention can be
administered as pharmaceutical compositions. Such
30 pharmaceutical compositions for use in accordance with the
present invention may be formulated in conventional manner
using one or more physiologically acceptable carriers or
excipients.

Thus, the compounds and their physiologically acceptable
35 salts and solvates may be formulated for administration by
inhalation or insufflation (either through the mouth or the
nose) or topical, oral, buccal, parenteral or rectal

administration.

The inhibiting compositions of the invention can further be administered in conjunction with standard dental techniques for the amelioration and/or prevention of
5 macrophage-mediated inflammatory damage associated with, for example, periodontal disease. The compositions can be administered in any form convenient for application to the gum and sub-gum area, in the form of a toothpaste in any form (gel, powder, foam) or a mouthwash. Further, the
10 compositions can be administered during periodontal surgical procedures, deep cleanings and the like.

For oral administration, the pharmaceutical compositions may take the form of, for example, tablets or capsules prepared by conventional means with pharmaceutically
15 acceptable excipients such as binding agents (e.g., pregelatinised maize starch, polyvinylpyrrolidone or hydroxypropyl methylcellulose); fillers (e.g., lactose, microcrystalline cellulose or calcium hydrogen phosphate); lubricants (e.g., magnesium stearate, talc or silica);
20 disintegrants (e.g., potato starch or sodium starch glycolate); or wetting agents (e.g., sodium lauryl sulphate). The tablets may be coated by methods well known in the art. Liquid preparations for oral administration may take the form of, for example, solutions, syrups or suspensions, or they
25 may be presented as a dry product for constitution with water or other suitable vehicle before use. Such liquid preparations may be prepared by conventional means with pharmaceutically acceptable additives such as suspending agents (e.g., sorbitol syrup, cellulose derivatives or
30 hydrogenated edible fats); emulsifying agents (e.g., lecithin or acacia); non-aqueous vehicles (e.g., almond oil, oily esters, ethyl alcohol or fractionated vegetable oils); and preservatives (e.g., methyl or propyl-p-hydroxybenzoates or sorbic acid). The preparations may also contain buffer
35 salts, flavoring, coloring and sweetening agents as appropriate.

Preparations for oral administration may be suitably
formulated to give controlled release of the active compound.

For buccal administration the compositions may take the
form of tablets or lozenges formulated in conventional
5 manner.

For administration by inhalation, the compounds for use
according to the present invention are conveniently delivered
in the form of an aerosol spray presentation from pressurized
packs or a nebulizer, with the use of a suitable propellant,
10 e.g., dichlorodifluoromethane, trichlorofluoromethane,
dichlorotetrafluoroethane, carbon dioxide or other suitable
gas. In the case of a pressurized aerosol the dosage unit
may be determined by providing a valve to deliver a metered
amount. Capsules and cartridges of, e.g., gelatin for use in
15 an inhaler or insufflator may be formulated containing a
powder mix of the compound and a suitable powder base such as
lactose or starch.

The compounds may be formulated for parenteral
administration by injection, e.g., by bolus injection or
20 continuous infusion. Formulations for injection may be
presented in unit dosage form, e.g., in ampoules or in multi-
dose containers, with an added preservative. The
compositions may take such forms as suspensions, solutions or
emulsions in oily or aqueous vehicles, and may contain
25 formulatory agents such as suspending, stabilizing and/or
dispersing agents. Alternatively, the active ingredient may
be in powder form for constitution with a suitable vehicle,
e.g., sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal
30 compositions such as suppositories or retention enemas, e.g.,
containing conventional suppository bases such as cocoa
butter or other glycerides.

In addition to the formulations described previously,
the compounds may also be formulated as a depot preparation.
35 Such long acting formulations may be administered by
implantation (for example subcutaneously or intramuscularly)
or by intramuscular injection. Thus, for example, the

- 27 -

compounds may be formulated with suitable polymeric or
hydrophobic materials (for example as an emulsion in an
acceptable oil) or ion exchange resins, or as sparingly
soluble derivatives, for example, as a sparingly soluble
5 salt.

The compositions may, if desired, be presented in a pack
or dispenser device which may contain one or more unit dosage
forms containing the active ingredient. The pack may for
example comprise metal or plastic foil, such as a blister
10 pack. The pack or dispenser device may be accompanied by
instructions for administration.

Toxicity and therapeutic efficacy of such compounds can
be determined by standard pharmaceutical procedures in cell
cultures or experimental animals, e.g., for determining the
15 LD_{50} (the dose lethal to 50% of the population) and the ED_{50}
(the dose therapeutically effective in 50% of the
population). The dose ratio between toxic and therapeutic
effects is the therapeutic index and it can be expressed as
the ratio LD_{50}/ED_{50}. Compounds which exhibit large therapeutic
20 indices are preferred. While compounds that exhibit toxic
side effects may be used, care should be taken to design a
delivery system that targets such compounds to the site of
affected tissue in order to minimize potential damage to
uninfected cells and, thereby, reduce side effects.

25 The data obtained from the cell culture assays and
animal studies can be used in formulating a range of dosage
for use in humans. The dosage of such compounds lies
preferably within a range of circulating concentrations that
include the ED_{50} with little or no toxicity. The dosage may
30 vary within this range depending upon the dosage form
employed and the route of administration utilized. For any
compound used in the method of the invention, the
therapeutically effective dose can be estimated initially
from cell culture assays. A dose may be formulated in animal
35 models to achieve a circulating plasma concentration range
that includes the IC_{50} (i.e., the concentration of the test
compound which achieves a half-maximal inhibition of

- 28 -

symptoms) as determined in cell culture. Such information
can be used to more accurately determine useful doses in
humans. Levels in plasma may be measured, for example, by
high performance liquid chromatography.

5

5.3. METHODS FOR AMELIORATION OF OSTEOSCLEROTIC BONE DISORDER-ASSOCIATED SYMPTOMS

Described herein are methods for the amelioration of
symptoms caused by osteosclerotic bone disorders, including
10 but not limited to PYCNO. In one embodiment of such methods,
the methods serve to increase the level of cathepsin K gene
and/or enzyme activity. In an additional embodiment, such
methods serve to increase both cathepsin K gene and/or enzyme
activity as well as increasing the level of at least a second
15 gene and/or enzyme activity involved in causing
osteosclerotic bone disorders. In a particular embodiment,
such methods increase the level of cathepsin K and cathepsin
S gene and/or enzyme activity.

As an example for purposes of clarity and not by way of
20 limitation, the description herein will frequently use as an
example an increase in cathepsin K gene and/or enzyme
activity.

With respect to an increase in cathepsin K enzyme
activity, enzyme replacement therapy can, for example, be
25 utilized. By following standard techniques, levels of
cathepsin K enzyme can be augmented to a concentration which
ameliorates symptoms of osteosclerotic bone disorders,
including PYCNO. Enzyme replacement, here, involves the
administration of exogenous enzyme, for example cathepsin K
30 enzyme, into a patient at a concentration sufficient to
ameliorate osteosclerotic disorder symptoms. Administration
can, for example, be intravenous. Alternatively, enzyme can
be administered directly to a site or sites of desired
action, in this case to bone tissue, especially to the site
35 of bone resorption, including the site of bone remodeling and
matrix degradation.

The cathepsin K enzyme utilized will include,
preferably, the proenzyme form of cathepsin K. In instances
in which the osteosclerotic disorder involves an inability to
successfully process the proenzyme into the active form, the
5 active form of cathepsin K is administered. In these
instances, site-specific, rather than systemic administration
is preferred. Dosage ranges necessary for amelioration of
osteosclerotic disorders can readily be determined by one of
skill in the art following, for example, techniques such as
10 those described, above, in Section 5.2.1. Such dosage ranges
can include, for example, an administration of enzyme at a
concentration ranging from about 0.1 μg/kg to about 10 mg/kg,
preferably from about 0.1 mg/kg to about 2 mg/kg.

With respect to an increase in the level of normal
15 cathepsin K gene expression, cathepsin K nucleic acid
sequences can be utilized for the amelioration of symptoms
caused by osteosclerotic bone disorders. Where the cause of
such a disorder is a defective cathepsin K gene, treatment
can be administered, for example, in the form of gene
20 replacement therapy. Specifically, one or more copies of a
normal cathepsin K gene or a portion of the cathepsin K gene
that directs the production of a cathepsin K gene product
exhibiting normal function, may be inserted into the
appropriate cells within a patient or animal subject, using
25 vectors which include, but are not limited to adenovirus,
adeno-associated virus, retrovirus and herpes virus vectors,
in addition to other particles that introduce DNA into cells,
such as liposomes.

Appropriate cells include cell types in which cathepsin
30 K gene expression is normally observed, especially osteoclast
cells. Administration techniques can involve, for example,
direct administration of cathepsin K gene sequences to the
site of the cells in which the cathepsin K gene sequences are
to be expressed, such as, for example, to the site of
35 osteoclasts within bone tissue, and, especially, the site of
bone remodeling, including the site of bone resorption.

Alternatively, targeted homologous recombination can be utilized to correct the defective endogenous cathepsin K gene in the appropriate tissue; e.g., osteoclast cells or osteoclast precursor cells. In non-human animals, targeted
5 homologous recombination can be used to correct the defect in ES cells in order to generate offspring with a corrected cathepsin K and, therefore, a corrected trait.

Additional methods which may be utilized to increase the overall level of cathepsin K gene expression and/or cathepsin
10 K activity include the introduction of appropriate cathepsin K-expressing cells, preferably autologous cells, into a patient at positions and in numbers which are sufficient to ameliorate the symptoms of osteosclerotic bone disorders, including PYCNO. Such cells may be either recombinant or
15 non-recombinant.

Among the cells which can be administered to increase the overall level of cathepsin K gene expression in a patient are normal cells, preferably osteoclast cells which express the cathepsin K gene, and/or osteoclast precursor cells which
20 give rise to cathepsin K-expressing osteoclast cells. Such cell-based gene therapy techniques are well known to those skilled in the art, see, e.g., Anderson, et al., U.S. Patent No. 5,399,349; Mulligan & Wilson, U.S. Patent No. 5,460,959.

In a particular embodiment, an autologous bone marrow
25 procedure can be utilized to isolate osteoclast precursor cells from a patient suffering from an osteosclerotic bone disorder or from an individual predisposed to such a disorder. The isolated cells can be modified so as to contain a normal cathepsin K gene and reintroduced into the
30 recipient so that osteoclast cells which express normal cathepsin K gene activity are produced in the patient, thereby ameliorating symptoms of the osteosclerotic disorder.

In another particular embodiment, the patient is deficient in both cathepsin K gene and/or enzyme activity as
35 well as in a second gene whose absence is involved in osteosclerotic bone disorders. In such an embodiment, the isolated osteoclast precursor cells are modified so as to

contain a normal cathepsin K gene and a normal second gene.
The modified isolated precursor cells are reintroduced into
the recipient so that osteoclast cells which express normal
cathepsin K gene activity and normal second gene activity are

5 produced in the patient, thereby ameliorating symptoms of the
osteosclerotic disorder. In a particular example of such an
embodiment, the second gene is a cathepsin S gene.

Because isolated bone marrow cells will contain
osteoclast precursor cells, bone marrow can be obtained using

10 standard techniques well known to those of skill in the art
and modified, as described above, with no further cell
separation or cell type purification.

In instances in which the cells utilized are to be
autologous cells, the bone marrow is obtained form the

15 patient to be treated. In instances in which the cells
utilized are heterologous cells, the cells are chosen
according to standard heterologous transplant criteria well
known to those of skill in the art.

Alternatively, standard separation techniques can be

20 used for the purification or partial purification of
osteoclast precursor cells. Such separation techniques are
generally based on the presence or absence of specific cell
surface markers, preferably transmembrane markers, and the
appropriate osteoclast precursor cell surface marker or

25 combination of markers which are unique to the osteoclast
precursor cells. Such a cell surface marker or combination
of markers are well known to those of skill in the art.

Purification techniques can take advantage of such a
cell surface marker or combination of markers by, for

30 example, utilizing antibodies which bind the specific
markers. These techniques can include, for example, flow
cytometry using a fluorescence activated cell sorter (FACS)
and specific fluorochromes, biotin-avidin or
biotin-streptavidin separations using biotin conjugated to

35 cell surface marker-specific antibodies and avidin or
streptavidin bound to a solid support such as affinity column

matrix or plastic surfaces or magnetic separations using —
antibody-coated magnetic beads.

A common technique for antibody based separation is the
use of flow cytometry such as by a fluorescence activated
5 cell sorter (FACS). Typically, separation by flow cytometry
is performed as follows. The suspended mixture of cells is
centrifuged and resuspended in media. Antibodies which are
conjugated to fluorochrome are added to allow the binding of
the antibodies to specific osteoclast precursor cell surface
10 markers. The cell mixture is then washed by one or more
centrifugation and resuspension steps. The mixture is run
through a FACS which separates the cells based on different
fluorescence characteristics. FACS systems are available in
varying levels of performance and ability, including multi-
15 color analysis. The osteoclast precursor cell can be
identified by a characteristic profile of forward and side
scatter which is influenced by size and granularity, as well
as by expression of cell surface markers.

Other separation techniques besides flow cytometry can
20 also provide fast separations. One such method is biotin-
avidin based separation by affinity chromatography.
Typically, such a technique is performed by incubating cells
with biotin-coupled antibodies to specific markers, followed
by passage through an avidin column. Biotin-antibody-cell
25 complexes bind to the column via the biotin-avidin
interaction, while other cells pass through the column. The
specificity of the biotin-avidin system is well suited for
rapid positive separation. Multiple passages can ensure
separation of a sufficient level of the osteoclast precursor
30 cell subpopulation of interest.

Upon isolation, and further separation if desired, if
cell viability and recovery are sufficient, cells can be
modified prior to introduction or reintroduction into the
patient by, for example, transformation with gene sequences
35 encoding a normal cathepsin K gene, and/or at least a second
gene encoding a second gene product involved in bone
resorption processes. Such modification is generally

contemplated where autologous cells are concerned. For
example, sequences can be introduced into the isolated
autologous cells which encode a normal cathepsin K and normal
cathepsin S gene products. Cell transformation and gene
5 expression procedures are well known to those of skill in the
art.

Cells, whether modified or not, can be cultured and
expanded *ex vivo* prior to administration to a patient.
Expansion can be accomplished via well known techniques
10 utilizing physiological buffers or culture media in the
presence of appropriate expansion factors such as
interleukins and other well known growth factors.

Cells are introduced or reintroduced into the patient
via standard techniques well known to those of skill in the
15 art. For example, cells can be washed, suspended in, for
example, buffered saline, and reintroduced into the patient
via intravenous administration. The osteoclast precursor
cells within the population of administered cells can rise to
a subpopulation of cells which will differentiate into
20 osteoclast cells. Such osteoclast cells will be capable of
expressing a normal cathepsin K gene and/or a normal second
gene, such as, for example, a normal cathepsin S gene, either
endogenous or recombinant, thereby ameliorating
osteosclerotic bone disorder symptoms.

25

    5.4. METHODS FOR IDENTIFICATION OF COMPOUNDS FOR AMELIORATION
         OF BONE RESORPTION DISORDER-ASSOCIATED SYMPTOMS

A variety of methods can be utilized for identifying
compounds capable of being used in ameliorating symptoms
30 caused by bone resorption disorders, including but not
limited to, osteoporosis, arthritides such as osteoarthritis,
and periodontal disease, and damage caused by macrophage-
mediated inflammatory processes.

The compounds identified via such methods are inhibitory
35 compounds capable of specifically inhibiting cathepsin K
activity and/or specifically inhibiting at least a second
activity involved in bone resorption disorders and/or

macrophage-mediated inflammatory damage processes. Such
activities and specificities are as described, above, in
e.g., Section 5.1. The identification methods described
herein comprise isolated enzyme-based assays, cell-based
5 assays and whole animal assays.

Enzyme-based assays can be utilized for rapid
identification of specific inhibitory compounds. In general,
such assays comprise contacting an enzyme to a model
substrate for the enzyme, in the presence of a test compound
10 for a time sufficient for the enzyme to act on the substrate.
The level of substrate acted on by the enzyme is measured to
determine the enzyme's level of activity. If the level of
enzymatic activity observed in the presence of that compound
is less then that observed in its absence, then that compound
15 is considered an inhibitor of the enzyme. In order to
determine whether the inhibitor is a specific inhibitor, the
test compound is assayed with other enzyme/enzyme substrate
pairs to determine whether it is inhibiting other enzymatic
activities as well. If the test compound inhibits enzymatic
20 activity of the first enzyme to a greater extent than it is
inhibits activity of other agents of interest, the test
compound is regarded as a specific inhibitor of the first
enzyme.

Taking cathepsin K as an example, such as enzyme-based
25 assay can include, first, contacting a model cysteine
protease substrate with cathepsin K in the presence of a test
compound for a time sufficient for cathepsin K to act on the
substrate. The level of hydrolyzed substrate is then
measured to determine cathepsin K enzymatic activity in the
30 presence of test compound. The test compound is also tested
in the presence of other individual cathepsin enzymes and the
model cysteine protease substrate and the enzymatic
activities of these individual additional cathepsins is also
determined. If the test compound reduces the enzymatic
35 activity of cathepsin K to a greater extent than it reduces
the enzymatic activity of the other cathepsins, the test
compound represents a cathepsin K specific inhibitor. Among

the other cathepsins which can be measured are, for example,
cathepsin S, B, L or H. Such an enzyme-based assay can also
be utilized to identify specific inhibitors of another of the
cathepsins which are measured. For example, a cathepsin S,
5 B, L or H specific inhibitor can be identified via such an
assay.

Such an enzyme-based assay can be expanded to further
test the specificity of the inhibitor to test, for example
whether it exhibits inhibiting action against other, non-
10 cysteine protease enzymes, including, but not limited to
serine proteases, metalloproteases and aspartyl proteases.

In cases where the inhibitor is contemplated for use in
inhibition of cathepsin K and an additional activity, an
inhibitor is also considered to be a specific inhibitor if it
15 reduces the level of cathepsin K and the additional activity,
for example, cathepsin S activity, to a greater extent than
it reduces the other activity levels tested. The enzymes
utilized in such assays will be in their active forms upon
contacting with substrate and test compound. Taking
20 cathepsin K as an example, a procathepsin K gene product can
be isolated from a cell expressing a cathepsin K gene using
standard cell expression, lysis and protein purification
techniques well known to those of skill in the art. The
isolated or partially isolated procathepsin K gene product
25 can then be activated into cathepsin K. Conversion of the
proform into the active enzyme form can be accomplished, for
example, by treatment with pepsin as described, for example,
in Bromme et al., 1996, J. Biol. Chem. 271:2126-2132, which
is incorporated herein by reference in its entirety.
30 Alternatively, procathepsin K can be heat activated in the
presence of cysteine, as described by Bossard et al., 1996,
J. Biol. Chem. 271:12517-12524, which is incorporated herein
by reference in its entirety.

Substrates utilized for such enzyme-based assays should
35 be ones whose cleavage, by hydrolysis, etc., by the enzyme of
interest can readily be measured. For example, in the case
of cathepsins, the substrates can be fluorogenic peptide

- 36 -

substrates whose cleavage can be measured by standard
fluorimetric methods. Among the cathepsin substrates which
can be utilized are methylcoumarylamide (MCA) peptides. A
substrate generally considered a model cysteine protease
5 substrate is Z-Phe-Arg-MCA, where Z is carboxybenzyl.
Cathepsin K favored substrates can also be utilized. For
example, MCA peptides of the general structure $Z-P_3-P_2-P_1-MCA$
(adapted from Bossard, M.J. et al., 1996, J. Biol. Chem.
271:12517-12524 and Bromme, D. et al., 1996, J. Biol. Chem.
10 271:2126-2132), where Z is carboxybenzyl, P_1 is preferably an
amino acid residue having a hydrophilic side chain and P_2 is
an amino acid having a small, hydrophobic side chain. In
general, amino acid residues with hydrophobic side chains at
the P_1 position are not favorable. In certain instances, such
15 substrates can also act as enzyme inhibitors.

Collagen and/or osteonectin can also be utilized as
cathepsin substrates in the identification methods of the
invention. Such substrates can be labeled, for example, by
utilizing standard techniques well known to those of skill in
20 the art, in order to follow and measure their degradation.

It is further noted that certain substrates favorable to
one cathepsin represent poor substrates for other cathepsin
enzymes, which can be utilized as one means for identifying
such cathepsin-specific inhibitors. For example, cathepsin K
25 and cathepsin S enzymes exhibit similar substrate
specificities. Both cathepsin K and cathepsin S, for
example, recognize Z-Leu-Arg-MCA preferentially to Z-Phe-Arg-
MCA, while cathepsin L exhibits an opposite preference. As
another example, Z-Arg-Arg-MCA represents a poor cathepsin K,
30 S and L substrate, but is readily recognized by cathepsin B.

Other differences between cathepsin K and other cysteine
proteases such as, for example, cathepsin B and L, can be
taken advantage of in identifying cathepsin K inhibitors.
For example, cathepsin K exhibits an increased stability at
35 neutral pH range. Further, cathepsin K exhibits an ability
to efficiently hydrolyze Type I collagen and elastin at pH
values above 6.0, and shows potent gelatinase activity in the

- 37 -

pH range of 5-7, while cathepsin L, for example, is limited
to a pH range of 4-5.5.

Among the test compounds which can be assayed for their
specific inhibitory properties are, for example, peptidyl
5 fluoromethyl ketones, vinyl sulfones, peptide aldehydes,
nitriles, α-ketocarbonyl compounds, including, for example,
α-diketones, α-keto esters, α-ketoamides, and α-ketoacids,
halomethyl ketones, diazomethyl ketones, (acyloxy)-methyl
ketones, ketomethylsulfonium salts and epoxysuccinyl
10 compounds.

Among the methods for the identification of inhibitors
as well as specific inhibitors, of cysteine protease enzymes
are, for example, techniques described in Riese, R.J. et al.,
1996, Immunity 4:357-366; Palmer, J.T. et al., 1995, J. Med.
15 Chem. 38:3193-3196; Bromme, D. et al., 1996, J. Biol. Chem.
271:2126-2132; Bossard, M.J. et al., 1996, J. Biol. Chem.
271:12517-12524; WO 95/23222; U.S. Patent No. 5,486,623; U.S.
Patent No. 5,525,623; U.S. Patent No. 5,158,936; and U.S.
Patent No. 5,374,623, all of which are incorporated herein by
20 reference in their entirety.

Compounds, including any of the compounds identified via
the enzyme-based assays described above, can be tested in
cell-based assays to test inhibitory capacity within the
cell. Such cell-based assays can be utilized to test a
25 number of features of a potential inhibitor, including, for
example, the compound's ability to enter the cell, its
cytoxicity, as well as its ability to act as an inhibitor
once inside the cell. Further, cell-based assays can
function to identify compounds which act more indirectly to
30 inhibit enzyme activity by one or more of the activities
listed, for example, in Section 5.1, above.

A typical cell-based assay involves contacting a cell
expressing the activity of interest with a test compound for
a time and measuring the inhibition of such an activity. For
35 measurements, for example, whole cells can be lysed according
to standard techniques and tested for the presence of various
enzyme activities. Among preferred cell-based assays is ones

such as that described in Riese et al. (Riese, R.J. et al.,
1996, Immunity 4:357-366, which is incorporated herein by
reference in its entirety).

Whole cell inhibitor assays can also be utilized. Such
5 assays serve to test inhibitor capacity in an in vivo
situation. Typically, such assays include administering to
an animal a test compound and measuring its inhibitory
effect. With respect to inhibitors of processes involved in
bone resorption disorders, assays for inhibitory capacity can
10 include, for example, standard measurements of hydroxyproline
levels in urine and bone density measurements.

Any of the compounds identified via the techniques
described herein can be formulated into pharmaceutical
compositions as described above and utilized as part of the
15 amelioration methods of the invention.

An inhibitor identified via the assay methods described
herein can directly affect activity, such as, in this case,
cathepsin K enzymatic activity, but can also alternatively
act to suppress or inhibit cathepsin K gene expression. For
20 example, inhibitor compounds can include, for example,
specific inhibitors of cathepsin K gene transcription and/or
translation. As discussed in Section 5.1.1, above, such gene
expression inhibitors can include, but are not limited to,
cathepsin K antisense and/or ribozyme inhibitors.

25 The specific inhibitors of a second activity involved in
the bone resorption process which can be identified via the
methods of the invention include, but are not limited to,
inhibitors of a gene product involved in the biosynthesis,
processing, activation, secretion, catalytic function, and/or
30 stability of cathepsin K or cathepsin K gene expression; a
gene product involved the availability or prior processing of
the cathepsin K substrate(s) for hydrolysis by cathepsin K;
another cathepsin, such as, for example, cathepsin S, L or B;
a protein involved in the generation of active osteoclasts
35 from their monocyte/macrophage progenitors; a protein
involved in the function of osteoclasts such as, for example,
in creating the acidic, sealed microenvironment of the

subosteoclastic space; a regulatory gene that controls the
tissue-specific upregulation of a protein involved in
cathepsin K biosynthesis, processing, secretion, catalytic
function, inhibition, or stability.

5

5.5. CATHEPSIN K-EXPRESSING AND CATHEPSIN K-DEFICIENT CELLS

For clarity of discussion, not limitation, the
generation of cell lines expressing cathepsin, or cathepsin
deficient cell lines, are described by way of example for the
10 cathepsin K gene. However, the principles may be analogously
applied to expression and generation of cell lines expressing
other cysteine cathepsins, metalloproteases, or other bone
degradative enzymes, or combinations, thereof, including
combinations comprising cathepsin K. Further, the principles
15 may also be analogously applied to expression and generation
of cell lines deficient for other cysteine cathepsins,
metalloproteases, or other bone degradative enzymes, or
combinations, thereof, including combinations comprising
cathepsin K deficient cells.
20   In accordance with the invention, cathepsin K nucleotide
sequences which encode cathepsin K, cathepsin K fusion
proteins or functional equivalents thereof may be used to
generate recombinant DNA molecules that direct the expression
of cathepsin K, or a functional equivalent thereof, in
25 appropriate host cells. The cathepsin nucleotide sequences,
which can include, for example, cDNA or genomic cathepsin
nucleotide sequences, may be engineered in order to alter the
cathepsin coding region for a variety of ends including but
not limited to alterations which modify processing and
30 expression of the gene. product. For example, mutations may
be introduced using techniques which are well known in the
art, e.g., site directed mutagenesis, to insert new
restriction sites, to alter glycosylation patterns,
phosphorylation, etc. Further, such cathepsin coding regions
35 can express pre-pro- forms, pro- forms and/or mature forms of
cathepsins.

In order to express a biologically active cathepsin K, the nucleotide sequence coding for cathepsin K, or a functional equivalent, is inserted into an appropriate expression vector, i.e., a vector which contains the

5 necessary elements for the transcription and translation of the inserted coding sequence. For example, among the regulatory sequences which may be utilized are upstream AP1 promoter sequences. It is to be recognized that the terms "express" or "expression" can refer to either expression or

10 overexpression relative to corresponding non-engineered cells.

In addition, cells or cell lines expressing cathepsin can, for example, be derived from transgenic animals, genetically engineered to overexpress cathepsin K, utilizing

15 standard techniques well known to those of skill in the art.

The cathepsin K gene products, as well as host cells or cell lines transformed or transfected with recombinant cathepsin K expression vectors can be used for a variety of purposes. These include screening and selection of drugs

20 that act to regulate the activity of cathepsin.

The term "cathepsin-deficient", as used herein, refers to cells, cell lines and/or animals which exhibit a lower level of functional cathepsin activity than corresponding cells, or cell lines or animals whose cells, contain two normal, wild

25 type copies of the cathepsin gene. Preferably, "cathepsin-deficient" refers to an absence of detectable functional cathepsin activity.

Cells and cell lines deficient in cathepsin activity can be derived from cathepsin knockout animals, utilizing

30 standard techniques well known to those of skill in the art. Such animals may be used to derive a cell line which may be used as an assay substrate in culture. While primary cultures may be utilized, the generation of continuous cell lines is preferred. For examples of techniques which may be

35 used to derive a continuous cell line from the transgenic animals, see Small et al., 1985, Mol. Cell Biol. 5:642-648. Such techniques for generating cells and cell lines can also

be utilized in the context of the transgenic and genetically
engineered animals described below.

With respect to cathepsin-K deficient cells, such cells
can, for example, also include cells taken from and cell
5 lines derived from patients exhibiting osteosclerotic
disorders, such as PYCNO, a disorder involving a cathepsin K
deficiency.

Additional cathepsin-deficient cells and cell lines can
be generated using well known recombinant DNA techniques such
10 as, for example, site-directed mutagenesis, to introduce
mutations into cathepsin gene sequences which will disrupt
cathepsin activity.

5.6. TRANSGENIC ANIMALS

15   Described in this Section are, first, animals expressing
one or more transgenic cathepsin gene sequences. Further,
animals deficient for one or more cathepsin genes are
described. For clarity of discussion, not limitation, the
generation of transgenic animals expressing or deficient for
20 cathepsin are described by way of example for the cathepsin K
gene. However, the principles may be analogously applied to
expression and generation of animals expressing other
cysteine cathepsins, metalloproteases, or other bone
degradative enzymes, or combinations, thereof, including
25 combinations comprising cathepsin K. Further, the principles
may also be analogously applied to expression and generation
of animals deficient for other cysteine cathepsins,
metalloproteases, or other bone degradative enzymes, or
combinations, thereof, including combinations comprising
30 cathepsin K deficiencies.

The present invention relates, for example, to the
generation of transgenic animals which contain a cathepsin
transgene. Such transgenic animals may serve as animal
models of bone readsorption disorders, including but not
35 limited to osteoporosis, arthritides, peridontal disease and
damage caused by macrophage-mediated inflammatory processes.
The transgenic animals of the invention can be utilized in

assay systems for identification of compounds capable of
inhibiting the activity of cathepsin, or cathepsin like
enzymes, thereby decreasing bone resorption.

In an embodiment of the invention, the role of cathepsin
5 K alone, or in combination with other cysteine cathepsins,
metalloproteases, or other bone degradative enzymes can be
investigated by genetically engineering transgenic animals to
express or overexpress one or more cathepsins or any of the
other bone degradative enzymes *in vivo*. The transgenic
10 animals may be genetically engineered to overexpress a single
cathepsin gene, multiple cathepsin genes, and/or other genes
encoding bone degradative enzymes.

Alternatively, transgenic animals may be engineered to
be deficient for one or more cathepsin activities. The term
15 "cathepsin-deficient" refers to animals which exhibit a lower
level of functional cathepsin activity than corresponding
animals whose cells, contain two normal, wild type copies of
the cathepsin gene. Preferably, "cathepsin-deficient" refers
to an absence of detectable functional cathepsin activity.
20   The transgenic animals of the invention can be any
species, including but not limited to mice, rats, rabbits,
guinea pigs, pigs, micro-pigs, and non-human primates, e.g.,
baboons, squirrel monkeys and chimpanzees. Any technique
known in the art may be used to introduce a transgene, such
25 as an inactivating or overexpressing gene sequence, into
animals to produce the founder lines of transgenic animals.
Such techniques include, but are not limited to pronuclear
microinjection (Hoppe, P.C. and Wagner, T.E., 1989, U.S. Pat.
No. 4,873,191); retrovirus mediated gene transfer into germ
30 lines (Van der Putten et al., 1985, Proc. Natl. Acad. Sci.,
USA 82:6148-6152); gene targeting in embryonic stem cells
(Thompson et al., 1989, Cell 56:313-321); electroporation of
embryos (Lo, 1983, Mol Cell. Biol. 3:1803-1814); and sperm-
mediated gene transfer (Lavitrano et al., 1989, Cell 57:717-
35 723); etc. For a review of such techniques, see Gordon,
1989, Transgenic Animals, Intl. Rev. Cytol. 115:171-229,
which is incorporated by reference herein in its entirety).

As listed above, standard embryonal stem cell (ES)
techniques can, for example, be utilized for generation of
cathepsin overexpressing, or cathepsin knockout and
cathepsin-deficient animals. ES cells can be obtained from
5 preimplantation embryos cultured in vitro (See, e.g., Evans,
M.J. et al., 1981, Nature 292:154-156; Bradley, .O. et al.,
1984, Nature 309:255-258; Gossler et al., 1986, Proc. Natl.
Acad. Sci. USA 83:9065-9069; Robertson et al., 1986, Nature
322:445-448; Wood, S.A. et al., 1993, Proc. Natl. Acad. Sci.
10 USA 90:4582-4584.) The introduced ES cells thereafter
colonize the embryo and contribute to the germ line of a
resulting chimeric animal (Jaenisch, R., 1988, Science
240:1468-1474).

Further, standard techniques such as, for example,
15 homologous recombination, coupled with cathepsin sequences,
including cathepsin K nucleotide sequences, can be utilized
to inactivate or alter any cathepsin genetic region desired.
A number of strategies can be utilized to detect or select
rate homologous recombinants. For example, PCR can be used
20 to screen pools of transformant cells for homologous
insertion, followed by screening of individual clones (Kim et
al., 1988, Nucl. Acids Res. 16:8887-8903; Kim et al., 1991,
Gene 103:227-233). Alternatively, a positive genetic
selection approach can be taken in which a marker gene is
25 constructed which will only be active if homologous insertion
occurs, allowing these recombinants to be selected directly
(Sedivy et al., 1989, Proc. Natl. Acad. Sci. USA 86:227-231).
Additionally, the positive-negative approach (PNS) method can
be utilized (Mansour et al., 1988, Nature 336:348-352;
30 Capecchi, 1989, Science 244:1288-1292; Capecchi, 1989, Trends
in Genet. 5:70-76). Utilizing the PNS method, nonhomologous
recombinants are selected against by using the Herpes Simplex
virus thymidine kinase (HSV-TK) gene and selecting against
its nonhomologous insertion with herpes drugs such as
35 gancyclovir or FIAU. By such counter-selection, the number
of homologous recombinants in the surviving transformants is
increased.

ES cells generated via techniques such as these, when
introduced into the germline of a nonhuman animal make
possible the generation of non-mosaic, _i.e._, non-chimeric
progeny. Such progeny will be referred to herein as founder
5 animals. Once the founder animals are produced, they may be
bred, inbred, outbred, or crossbred to produce colonies of
the particular animal.

Cathepsin transgenic animals can be generated using
cathepsin nucleotide sequences, for example cathepsin K
10 nucleotide sequences, which are well known to those of skill
in the art and/or which can routinely be isolated utilizing
standard molecular biological techniques. Such cathepsin
sequences can encode pre-pro forms, pro- forms, as well as
mature forms of cathepsins. Further, such cathepsin
15 sequences can include genomic sequences, cDNA sequences or
hybrids thereof.

The cathepsin nucleotide sequences that can be utilized
may be derived from any species, including but not limited to
mice, rats, rabbits, guinea pigs, pigs, non-human primates,
20 baboons, squirrel monkeys and chimpanzees. Such sequences
can be routinely isolated by utilizing standard molecular
techniques and human cathepsin K nucleotide sequences as
probes and/or as PCR primers, as discussed below. In
addition, nucleotide sequences encoding cathepsins other than
25 cathepsin K, such as cathepsin S or the cysteine cathepsins,
metalloproteases, or other bone degradative enzymes can be
isolated using standard molecular biological techniques and
utilized to generate transgenic animals.

A representative cathepsin-deficient, or "knockout"
30 animal is a mouse cathepsin-deficient animal. Knockout mice
are well known to those of skill in the art. See, for
example, Horinouchi, K. et al., 1995, Nature Genetics 10:288-
293; and Otterbach, B. & Stoffel, W., 1995, Cell 81:1053-
1061, both of which are incorporated herein by reference in
35 their entirety.

To accomplish cathepsin gene disruptions, the technique
of site-directed inactivation via gene targeting (Thomas,

- 45 -

K.R. and Capecchi, M.R., 1987, Cell 51:503-512) and review in
Frohman et al., 1989, Cell 56:145-147; Cappecchi, 1989,
Trends in Genet. 5:70-76; Barribault et al., 1989, Mol. Biol.
Med. 6:481-492; Wagner, 1990, EMBO J. 9:3025-3032; and
5 Bradley et al., 1992, Bio/Technology.

An inactive allele of the cathepsin K gene can be
generated by targeted mutagenesis using standard procedures
of combined positive and negative selection for homologous
recombination in embryonic stem (ES) cells. Cathepsin K
10 genomic clones can be isolated, for example, from a 129/sv
mouse genomic library, which is isogenic with the W9.5 line
of ES cells to be used for gene targeting, as demonstrated in
the Example presented in Section 9, below. The null
targeting vector can be constructed containing homologous
15 sequences flanking both 5' and 3' sides of a deletion of the
first coding exon (exon 2), including the translational
initiation codon, and other essential coding sequences of the
gene. The vector carries a resistance marker, e.g., a
neomycin resistance marker (Neo) for positive selection and a
20 negative marker, e.g., a thymidine kinase (TK) marker, for
negative selection. Vectors can be utilized which are
analogous to previously reported targeting vectors,
successfully used for generating knock-out mice for other
genes, e.g., for Niemann-Pick Disease, NMDA receptor and
25 thyroid hormone receptor.

Once such a cathepsin-deficient mouse is generated, it
would be possible to also generate doubly or multiply
cathepsin-deficient animals by, for example, beginning with
cells derived from the cathepsin K animal and repeating the
30 above procedures utilizing a second cathepsin, for example,
cathepsin S, gene sequence for inactivation. Additionally,
appropriate genetic crosses between animals deficient for
different individual genes can result in progeny deficient
for more than one gene of interest.

35   To create a transgenic animal overexpressing cathepsin
K, genomic DNA containing the cathepsin K gene can be
transferred into the genome of the transgenic animal using

the techniques detailed below. The transgenic sequence may
be derived from the same or a different species than the
animal itself.

For example, transgenic animals can include mice which
5 overexpress human cathepsin K transgenic sequences and/or
overexpress mouse transgenic cathepsin K sequences. Such
sequences can comprise genomic DNA, cDNA or hybrids thereof.
Human genomic clones containing the cathepsin K gene can be
isolated by screening a human genomic library by polymerase
10 chain reaction, using primers to exons 3 and 4, as
demonstrated in the Example presented, below, in Section 8.
A similar approach may also be utilized to isolate cathepsin
K clones from other species using a genomic library derived
from that species. The isolation of murine cathepsin K
15 genomic sequences is demonstrated, for example, in the
Example presented in Section 9, below. The genomic clones
may be assayed for the presence of both the upstream 5'
flanking region and the 3' untranslated region indicating the
presence of the entire, for example, human or murine
20 cathepsin K gene. In addition the clones may be tested for
the presence of cathepsin S which is closely linked to
cathepsin K. In an additional embodiment of the invention,
genomic DNA encoding other bone degradative enzymes, in
addition to cathepsin K, such as metalloproteases, can be
25 cotransferred with cathepsin genomic DNA resulting in
transgenic animals overexpressing multiple enzymes involved
in bone regeneration.

Transgenic animals overexpressing a cathepsin
nucleotide sequence, a metalloprotease nucleotide sequence,
30 and/or a nucleotide sequence encoding other bone degradative
enzymes, may also be generated using vectors containing said
nucleotide sequences under the control of suitable expression
control elements. The expression control elements to be used
will vary depending on the desired strength or specificity of
35 expression. As used herein, expression control elements
include but are not limited to inducible and non-inducible
promoters, enhancers, operators and other elements known to

those skilled in the art that drive and regulate expression.
Such expression control elements include, but are not limited
to, viral control sequences such as the cytomegalovirus hCMV
immediate early gene, the early or late promoter of SV40, or
5 retroviral LTR sequences. Other expression control elements
include the lac system, the trp system and promoters,
including, but not limited to, promoters which contain AP1
regulatory sequences, derived from the mammalian genome. In
instances where cell-type specific expression is required,
10 the regulatory sequences required for such cell-type specific
activation will depend upon the particular cell type of
interest, and will be apparent to those of skill in the art.
In the case of cDNA sequences, standard "mini gene"
techniques can be applied for successful expression and
15 overexpression of transgenic sequences.

     Taking one example of the generation of transgenic
animals, specifically transgenic mice, DNA encoding cathepsin
K, and or cathepsin S, can be injected into inbred FVB
zygotes as described by Gorden et al. (1980, Proc. Natl.
20 Acad.Sci. USA 77:7380-7384). Genomic or vector DNA can be
electroporated into W9.5 ES cells (male-derived), which can
then be cultured and selected on feeder layers of mouse
embryonic fibroblasts derived from transgenic mice expressing
a Neo gene. G418 (350 mg/ml; for gain of Neo) and
25 ganciclovir (2 mM; for loss of TK) can be added to the
culture medium to select for resistant ES cell colonies that
have undergone homologous recombination at the URO-D gene.
Recombinants can then be identified by screening genomic DNA
from ES cell colonies by Southern blot hybridization
30 analysis. Correctly targeted ES cell clones, which also
carry a normal complement of 40 chromosomes, can be used to
derive mice carrying the mutation. ES cells can be micro-
injected into blastocysts at 3.5 days post-coitum obtained
from C57BL/6J mice, and blastocysts will be re-implanted into
35 pseudopregnant female mice, which serve as foster mothers.
Chimeric progeny derived largely from the ES cells are
identified by a high proportion of agouti coat color (the

color of the 129/sv strain of origin of the ES cells) against
the black coat color derived from the C57BL/6J host
blastocyst. Male chimeric progeny are tested for germline
transmission of the mutation by breeding with C57BL/6J
5 females. Agouti progeny derived from these crosses are
expected to be heterozygous for the mutation, which will be
confirmed by Southern blot analysis. These F1 heterozygous
·progeny are then inter-bred to generate F2 litters containing
progeny of all three genotypes (wild type, heterozygous and
10 homozygous mutants) for phenotypic analyses.

        Transgenic mice overexpressing cathepsin K can be
screened by PCR using primers from the 5' flanking region and
the 3' untranslated region (UTR) to demonstrate the presence
of the entire genomic cathepsin K gene. Southern blot
15 analysis may also be performed using cathepsin DNA as a probe
· to detect the presence of the cathepsin transgene. Transgene
copy number may be estimated by quantitative analysis from
Southern blot autoradiography. Message levels for cathepsin
transgene expression may be determined using Northern
20 analysis from total RNA derived from tissue known to express
cathepsin. Similarly, immunoblot analysis may be performed
using whole cell lysates derived from tissues know to have
detectable quantities of cathepsin.

25 5.6.   BONE DENSITY AND CATHEPSIN K POLYMORPHISMS/MUTATIONS

        It is contemplated that polymorphisms and/or mutations
within the cathepsin K gene can affect bone density in
individuals carrying alleles containing such polymorphisms
and/or mutations. Methods are presented herein for the
30 detection of bone density affecting polymorphisms within the
cathepsin K gene.

        In general, such methods comprise the correlation of
specific polymorphisms and/or mutations of classes of such
polymorphisms and/or mutations with a specific bone density
35 phenotypic range. By detecting individuals exhibiting or
predisposed to bone density-related disorders, appropriate
therapeutic or prophylactic treatments can be prescribed.

Further, by determining the expected severity of the bone
density disorder generally associated with a specific
genotype, more treatments can be administered which are more
commensurate with the severity or expected severity of the
5 bone density disorder.

First, bone density measurements can be performed on an
individual to be tested, for example, an individual in a risk
group for a bone density disorder such as osteoporosis. Such
bone density measurements are well known to those of skill in
10 the art. For example, dual energy X-ray absorptimetry (DEXA)
measurements can be performed to assay bone mineral density.
See, e.g., Sim, L.H. and Doorn, T., 1995, Austral. Phys. Eng.
Sci. Med 18:65-80.

Next, cathepsin K polymorphisms and/or mutations can be
15 detected in the individual being tested, using standard
techniques which are, as described below, well known to those
of skill in the art. Upon measuring both bone density and
cathepsin K polymorphisms, a correlation between bone density
and the specific cathepsin K allele(s) present in the
20 individual can be assessed and compared. Within a population
of individuals, it can be possible to correlate a given bone
density with a given cathepsin K polymorphism. In certain
instances, the presence of a given polymorphism may so
tightly correlate with a given bone density that it may
25 become possible to carry out only the cathepsin K
polymorphism analysis, in the absence of the bone density
measurements.

Mutations within the cathepsin K gene can be detected by
utilizing a number of techniques. Nucleic acid from any
30 nucleated cell can be used as the starting point for such
assay techniques, and may be isolated according to standard
nucleic acid preparation procedures which are well known to
those of skill in the art.

DNA may be used in hybridization or amplification assays
35 of biological samples to detect abnormalities involving
cathepsin K gene structure, including point mutations,
insertions, deletions and chromosomal rearrangements. Such

assays may include, but are not limited to, Southern
analyses, single stranded conformational polymorphism
analyses (SSCP), and PCR analyses, all of which are well
known to those of skill in the art. The cathepsin K gene
5 sequences which can be used can include, for example,
cathepsin K cDNA (see, e.g., Shi, G.P. et al., 1995, FEBS
Lett. 357:129-134). Further, the cathepsin K sequences which
can be utilized can include the cathepsin K genomic sequences
disclosed herein. Still further, the cathepsin K genomic
10 structure, presented here for the first time, can be utilized
in analyzing the cathepsin K polymorphisms and/or mutations
present in a given individual to be tested.

6.   EXAMPLE: LOSS OF CATHEPSIN K ACTIVITY IS THE CAUSATIVE
15   AGENT OF PYCNODYSOSTOSIS
      The Example presented herein demonstrates the discovery
for the first time, of an agent which degrades bone matrix *in
vivo*. Specifically, this Example demonstrates that loss of
cathepsin K activity is responsible for pycnodysostosis,
20 (PYCNO) an autosomal recessive osteochondrodysplasia
characterized by osteosclerosis and short stature. The
findings thus define cathepsin K as a major protease involved
in matrix degradation, as exemplified here by osteoclast-
mediated matrix degradation associated with bone resorption.

25
      6.1 MATERIALS AND METHODS
      Cathepsin S exons 2-5 were amplified from genomic DNA
and sequenced by cycle sequencing using an ABI 377 sequencer.
For analysis of the remaining coding region, total
30 lymphoblast RNA was extracted (RNazol, Tel Test) and first-
strand CDNA was synthesized using an oligo dT primer (Gibco
BRL). Cathepsin S CDNA nucleotides 631-1231 were amplified
using gene-specific primers. Sequence comparisons were
carried out with the AutoAssembler software package (Applied
35 Biosystems)
      The R113W substitution in the cathepsin S gene was
assessed in genomic DNA by amplification of a 150 bp product

containing exon 4 and the adjacent intron 4 sequence using
the antisense (5'-GCATTTAAAGAGCTCTACCTAGGGTT-3') and the
sense (5'-CAGAGAAATATCACATATAAGTCAAACCCCAAT-3') primers, the
latter introduced a T-to-C change (underlined) that created a
5 *MunI* restriction site in the allele with the sequence
variation. The amplified product was isolated (Qiagen),
digested and analyzed on a 2% agarose gel.

   Cathepsin S enzymatic activity was assayed in
lymphoblast lysates fluorometricly using the substrate Z-Val-
10 Val-Arg-NHMec (Bachen Feinchemikalien AG, Bubendorf,
Switzerland) following inactivation of cathepsins B, L, and H
at 40° C (Kirschke, H. and Wiederanders, B., 1994, *Methods in
Enzymology*: 244:500, Academic Press, New York). The control
lysosomal enzyme, α-galactosidase A, was determined
15 fluorometrically (D. Bishop, F. and Desnick, R.J. 1982, *J.
Biol. Chem.* 256, 1307).

   Cathepsin K cDNA was amplified by RT-PCR from PYCNO
lymphoblast total RNA. For PCR amplification, cathepsin K
sense and antisense primers were synthesized corresponding to
20 the 5' and 3' untranslated regions (5'-GCCGCAATCCCGATGGA-3'
and 5'-CCTTGAGGATATTGAAGGGAACTTAG-3', respectively). Nested
sense and antisense PCR primers (5'-CCCCTGATGGTGTGCCCCA-3'
and 5'CCCTTCCAAAGTGCATCGTTACACT-3', respectively) were used
to reamplify the cathepsin K cDNA from the initial PCR
25 product. Products were isolated and cycle sequenced in both
orientations.

   The putative cathepsin K mutation in the affected
Israeli Arab PYCNO patients was assayed in family members and
in unrelated normal Arab individuals by amplifying
30 nucleotides 998-1170 from genomic DNA using flanking primers
(5'GGGGAGAAAACTGGGGAAACA-3' and the antisense primer from the
nested PCR) and then digesting with *HinfI* (New England
Biolabs).

35   6.2   RESULTS
   In an initial attempt to identify the PYCNO gene, a
genome-wide search for the PYCNO locus was undertaken in a

large, consanguineous Israeli Arab family with 16 affected
relatives (Gelb, B. D. et al., 1995, Nature Genet. 10:235).
A genomic region that was homozygous-by-descent for all
affected individuals was localized to the pericentric region
5 of chromosome I with a lod score of 11.72 at *D1S498*.
Ancestral recombinant events localized the PYCNO region to 4
cM from *D1S442* to *D1S305* (FIG. 1).

Independent linkage analysis of a large Mexican PYCNO
family confirmed linkage to the pericentric region of
10 chromosome I and excluded the macrophage colony stimulating
factor, *CSF1*, which is defective in the *op/op* mouse model of
osteopetrosis (Polymeropoulos, M. H. et al., 1995, Nature
Genet. 10:238). Recently, the PYCNO locus was narrowed to a
region of about 2 cM at chromosome 1q21 from *D1S2344* to
15 *D1S2343/D1S2345* by analyzing the Israeli Arab family with new
markers (Gelb, B. D. et al., 1996, Hum. Genet. 98:141-144;
FIG. 1). This refinement excluded the interleukin-6 receptor
gene which previously had been proposed as a candidate gene
(Gelb, B. D. et al., 1995, Nature Genet. 10:235), but
20 included, for example, *MCL1*, a *Bcl2* homolog relevant to
monocyte/macrophage differentiation.

An additional putative candidate gene was lysosomal
cathepsin S, a cysteine protease with expression in
osteoclasts which had been assigned to chromosome band 1q21
25 by FISH analysis (Shi, G.-P., et al., 1994, J. Biol. Chem.
269:11530). To determine whether cathepsin S mapped to the
PYCNO critical region, a cathepsin S sequence tagged site
(STS) was developed using the published 5'-flanking region
sequence (Shi, G.-P., et al., 1994, J. Biol. Chem.
30 269:11530), and STS content mapping was performed with a
yeast artificial chromosome (YAC) contig spanning the PYCNO
region utilizing standard techniques. The cathepsin S STS
was amplified from two overlapping CEPH megaYACs, 978e4 and
947e1, and was mapped centromeric to *D1S498* in the PYCNO
35 critical region (FIG. 1).

Sequencing of the cathepsin S coding region in affected
members from the Israeli Arab family and from a

consanguineous Moroccan Arab family by exon amplification
from genomic DNA and by RT-PCR revealed a C-to-T transversion
of a CpG dinucleotide at position 343 in the cDNA in both
families, predicting an R113W substitution in the
5 propeptide,near the putative cleavage site at I114 (Shi, G.P.
et al., 1992, J. Biol. Chem. 267:7258).

Analysis of 18 unrelated normal Israeli Arab individuals
living near the Israeli Arab PYCNO family demonstrated that
12 of 36 alleles were positive for R113W, consistent with it
10 being a sequence polymorphism. In addition, cathepsin S
activities in lymphoblasts from two Israeli Arab PYCNO
patients and three normal controls were not significantly
different (210 Mu/U α-gal A ± 100 vs 340 ±40, respectively)
Thus, cathepsin S was ruled out as a candidate gene for
15 PYCNO.

Although cathepsin S was ruled out by these studies as a
candidate gene for PYCNO it was determined, for the first
time, that another cathepsin cysteine protease, cathepsin K,
also mapped to the critical 1q21 region. This determination
20 was made via chromosomal fluorescence-labeled in situ
hybridization (FISH), using a labeled cathepsin K primer and
standard techniques, and also by an amplification study.

For the amplification study, a cathepsin K STS,
developed from the published 3' cathepsin K untranslated
25 region (Shi, G.-P. et al., 1995, FEBS Letters 357:129;
Bromme, D. & Okamoto, K., 1995, Biol. Chem. Hoppe-Seyler
376:37; Tezuka, K., et al., 1994, J. Biol. Chem. 269:1106;
Inaoka, T., et al., 1995, Biochem. Bioshys. Res. Commun.
206:89), was amplified from the same two overlapping YACs in
30 the PYCNO critical reaion that contained the cathepsin S STS.
RT-PCR amplification and sequencing of the cathepsin K
transcript from lymphoblast total RNA derived from two
Israeli Arab PYCNO patients revealed an A-to-G transition at
cDNA nucleotide 1095 (Shi, G.-P. et al., 1995, FEBS Letters
35 357:129; Bromme, D. & Okamoto, K., 1995, Biol. Chem. Hoppe-
Seyler 376:37; Tezuka, K., et al., 1994, J. Biol. Chem.
269:1106; Inaoka, T., et al., 1995, Biochem. Bioshys. Res.

Commun. 206:89.), which predicted the substitution of the
termination codon by a tryptophan residue (X330W) and the
elongation of the carboxy-terminus by 19 additional amino
acids (FIG. 2).

5   Evaluation of the X330W allele in the entire Israeli
Arab PYCNO family and 43 unrelated normal Arab control
individuals revealed that it cosegregated with disease in the
PYCNO family and was not present in any of the 86 Arab
control alleles. Further evidence that cathepsin K mutations
10 caused PYCNO was the finding of five additional mutations,
yielding a total of eight cathepsin K mutations within twelve
PYCNO families (FIG. 2). For example, two affected Moroccan
Arab siblings had a missense mutation, a G-to-C transversion
at nucleotide 541, predicting a G146R substitution. Further,
15 an American Hispanic patient with non-consanguineous parents
was found to be heterozygous for markers across the PYCNO
critical region. Sequence analysis demonstrated
heteroallelism for the G146R mutation and a C-to-T transition
of a CpG dinucleotide at nucleotide 826, predicting an R241X
20 nonsense mutation. Restriction analysis of amplified segments
from genomic DNA with *Banl* for G146R and *Aval* for R241X
confirmed the RT-PCR results. In addition, a nonsense
mutation, K52X, which predicts a severely truncated protein,
and E79G, a missense mutation, were both found in a
25 heterallelic form in a patient from a Utah family. Further,
the R241X and the missense mutation Y212C was a found in
patients from a Spanish PYCNO family. The Spanish family
contained two PYCNO patients from different generations, with
the older patient being homoallelic and the younger
30 heterallelic. Two additional missense mutations, A277E and R
321G, were found in families as indicated in FIG. 2. See
Fig. 2 for a complete summary or which alleles have been
found in the particular families.

       To assess the effects of the putative Israeli Arab
35 cathepsin K mutation on enzyme activity, the X330W mutation
was introduced by site-directed mutagenesis into the full-
length cDNA. The normal and X330W cathepsin K alleles were

transiently expressed in 293 cells which do not have
detectable cathepsin K protein (FIG. 3A). Cathepsin K
protein was observed by immunoblotting in the whole cell
lysates of cells transfected with the nominal allele.   In
5 contrast, no detectable protein was observed in lysates from
cells transfected with the X330W construct. Northern
analysis of total RNA from transfected cells revealed
comparable steady-state transcript levels between normal and
mutant for both cathepsin K and the co-transfected
10 adenovirus-associated I RNA gene (FIG. 3B).

     The cathepsin K X330W mutation altered the stop codon
and predicted an elongated polypeptide. Such mutations are
rare, the best characterized being Hb Constant Spring which
results in an unstable transcript, thereby decreasing α
15 globin chain synthesis (Clegg, B. et al., 1971; Nature
234:337; Rousseau, F. 1995, Nature Genet. 10:11-12;
Liebhaber, S.A. & Kan, Y.W. 1981, J. Clin. Invest. 68:439;
Hunt, D, M. et al., 1982, Br. J. Haematol. 51:405-413). In
contrast, the X330W cathepsin K polypeptide was not detected
20 in transfected cells despite normal transcript levels,
suggesting protein instability. The R241X mutation predicts
truncation of the polypeptide at residue 241, and the loss of
89 amino acids including the completely conserved H276 and
N296 residues of the active site (Shi, G.-P. et al., 1995,
25 FEBS Letters 357:129; Bromme, D. & Okamoto, K., 1995, Biol.
Chem. Hoppe-Seyler 376:37; Tezuka, K., et al., 1994, J. Biol.
Chem. 269:1106; Inaoka, T., et al., 1995, Biochem. Bioshys.
Res. Commun. 206:89.). In addition, truncation mutations are
known to cause transcript instability and degradation and
30 also can result in exon skipping (Dietz, H. C. et al., 1993,
Science 259:680), while the K52X mutation predicts an even
more severely truncated protein. Thus, the R241X and K52X
mutation, are presumably null for cathepsin K activity.
Further, each of the missense mutations cause significant
35 changes in amino acid residue types at the affected
positions, which can affect activity.

Little, if any, cathepsin K activity is detectable in
lymphoblasts or fibroblasts, limiting the laboratory
diagnosis of RT-PCR to DNA-based methods. As shown by these
studies, specific cathepsin K mutations causing PYCNO can be
5 detected by RTPCR and sequencing of transcripts from
lymphoblasts. The cathepsin K gene organization shown in
FIG. 2 and described below, in Section 8, will facilitate
direct mutation analysis from genomic DNA

   That PYCNO results from mutations in the cathepsin K
10 gene is notable. This pure skeletal dysplasia is now
classified as a lysosomal disease, resulting from the
deficient activity of a lysosomal protease. Most lysosomal
enzymes are constitutively expressed in all cell types
(except erythrocytes) and lysosomal disease phenotypes result
15 from the accumulation of their substrate(s) in various cells
and tissues. For example, the mucopolysaccharidoses and most
glycoproteinoses each has diverse manifestations such as
dysostosis multiplex, mental retardation, corneal dystrophy,
and hepatosplenomegaly. In contrast, the PYCNO phenotype,
20 which is restricted to bone, results from the deficiency of
an enzyme required for bone resorption by osteoclasts with
minimal storage of its substrate. The cell-specific high
expression of cathepsin K is unique among known lysosomal
enzymes defining PYCNO as a lysosomal disorder due to
25 defective tissue-specific expression.

   Recently, cysteine proteases have been implicated in
bone resorption and remodeling based on the immunohistologic
localization of cathepsins B and L to the subosteoclastic
space (Blair, H. C. et al., 1993, Clin. Orthop. 294:7; Goto,
30 T., 1994, Histochemistry 101:33), and by the *in vitro*
inhibition of bone matrix degradation using non-specific
cysteine cathepsin inhibitors (Debari, K. et al, 1995,
Calcif. Tissue Int. 56:566; Everts, V. et al., 1988, Calcif.
Tissue Int. 43:172; Everts, V. et al., 1992, J. Cell.
35 Physiol. 150:221; Kakegawa, H. et al., 1993, FEBS Letters
321:247). The finding that cathepsin K deficiency causes
PYCNO, taken together with the knowledge that cathepsin K is

the only cysteine protease highly expressed in osteoclasts
(Shi, G.-P. et al., 1995, FEBS Letters 357:129; Bromme, D. &
Okamoto, K., 1995, Biol. Chem. Hoppe-Seyler 376:37; Tezuka,
K., et al., 1994, J. Biol. Chem. 269:1106; Inaoka, T., et
5 al., 1995, Biochem. Bioshys. Res. Commun. 206:89) and that it
has the highest type I collagenolytic, elastinolytic, and
gelatinolytic activities of cysteine proteases (Bromme, D.
et. al., 1996, J. Biol. Chem. 271:2126), indicates that
cathepsin K is the major protease in bone matrix resorption.
10

7.   EXAMPLE: CATHEPSIN K MUTATIONS INDICATE PRESENCE OF
                 COMMON NON-CATHEPSIN PYCNO PREDISPOSING
                 MUTATIONS

     The Example presented in Section 6, above, demonstrated
the definitive identification of the loss of cathepsin K
15 activity to be the cause of the bone disorder, PYCNO. The
Example presented in this Section presents surprising data
that indicate the presence of a second common polymorphism or
mutation in a second gene which leads to a predisposition for
PYCNO.
20
     First, in a study of four presumably related families on
Madeira Island, Portugal, a small island with a population of
approximately 250,000 at least three separate cathepsin K
mutations were found in PYCNO individuals, PYCNO being a
disorder with a incidence of less than one in a million.
25 This finding suggested that a second gene locus with a more
common, even frequent, mutation contributed to the deficient
bone resorption causing PYCNO among these Portuguese
patients. Thus, five Portuguese PYCNO patients had four
different cathepsin K mutations. This is remarkable,
30 particularly since four of the families were individually
consanguineous and their affected members were homoallelic
for each of the four mutations and any closely linked genes.
     The concept of a predisposing polymorphism or mutation
in a second gene is supported by two findings. First, a
35 total of 16 presumably unrelated families with PYCNO have
been analyzed for cathepsin K mutations. In each family, all

- 58 -

affected members within families inherited the same cathepsin
K mutation inherited as an autosomal recessive trait. Of
these 16 families whose collection was based mostly on
previous case reports, nine or about 60% were of Hispanic
5 ancestry, suggesting a common predisposing polymorphism or
mutant gene, at least in the Hispanic patients. Also, the
family reported by Polymeropoulos, supra., was of Mexican
ancestry. Other such predisposing polymorphisms may occur in
other ethnic, demographic, or inbred groups. Such a
10 predisposing gene would cause PYCNO when cathepsin K activity
was also deficient.

    The second finding is based on an analysis of a large
Israeli Arab family, referred to above, in Section 6. PYCNO
inheritance in this family tracks perfectly with the
15 cathepsin K genetic region, suggesting that the second gene
might lie in the same immediate region of chromosome position
1q21. As discussed in Section 6, above, this region had
already been shown to contain cathepsin S and K. The region
could further contain additional cathepsins or other relevant
20 genes.

    The analysis revealed a mutation (R113W) in the
cathepsin S gene, another cysteine protease gene located
adjacent (<150 kb) to the cathepsin K gene on chromosome band
1q21. As discussed above, in Section 6, both the cathepsin K
25 and S genes, mapped to the PYCNO critical region between
D1S2344 and D1S2343/D1S2345. Further, both genes were found
to be present on the same human PAC (P1 artificial
chromosome; clone 74e16). The cathepsin S mutation was
inherited in the homozygous condition by all affected members
30 of the Israeli Arab family. It should be noted that the
R113W mutation may be involved in the processing of the
cathepsin S polypeptide to a functionally active form. It
should further be noted that the cathepsin S mutation was
common, occurring in about one-third of the Arab population
35 living in the same area of Israel as the affected family.
Thus, the expression of this disease may require alteration

in at least two genes, perhaps adjacent, as in the case of
cathepsins K and S in the Israeli Arab family.

This finding bears directly on the treatment of
disorders due to excessive bone resorption. If inhibitors of
5 cathepsin K are proposed to treat osteoporosis, the effects
of such interventions may depend on the nature of the
polymorphism or mutation in the second (or additional)
predisposing gene(s). For instance, if there is a second
cathepsin with overlapping function in bone resorption, an
10 inhibitor of cathepsin K the minimal affects on this second
cathepsin might be ineffective, whether the two cathepsins
work independently, dependently, or synergistically. Thus,
the presence of a second, presumably closely-linked locus
involved in the synthesis, processing, function, or transport
15 of cathepsin K, or directly in the bone resorption process,
is highly relevant to interventions in the bone resorption
process.

It is notable that another genetic disorder (limb-girdle
muscular dystrophy Type 2A) caused by the deficiency of a
20 neutral cysteine protease, calpain 3, was recently identified
in which presumably related families on a small island had
different mutations (Richard, I et al., 1995, Cell 81:27-40).
Calpain 3 belongs to a family of calpains, analogous to the
cathepsin family. The finding of multiple mutations in
25 Calpain 3 suggested to Richard et al. that a "digenic" model
in which only in the presence of specific alleles at a
permissive second locus (e.g., a compensatory, partially
redundant, regulatory, or modifier gene) will there be
expression of calpain mutations. Since one would need
30 mutations at both loci to be affected, the disease prevalence
would remain low. By analogy, it is possible that a similar
situation holds for cathepsin K, that another locus, perhaps
a cathepsin like S, is required for expression of the
disease. Therefore, it may require inhibitors of both genes
35 or enzymes to prevent or inhibit bone resorption.

8.   EXAMPLE: CATHEPSIN K GENOMIC STRUCTURE

In this Example, the genomic organization of the human cathepsin gene is presented, including the exon and intron sizes, the exon/intron boundary sequences, the transcription initiation site and the 5' flanking region. In addition, the
5 chromosomal location of the human gene was determined by fluorescence in situ hybridization (FISH). The isolation and dissection of the cathepsin K genomic structure is especially advantageous, given the fact that cathepsin K cDNA-based analyses are quite difficult utilizing standard techniques.
10 Specifically, the lymphoblast blood cells usually isolated for studies such as diagnostic analyses are virtually devoid of cathepsin K RNA, making these standard cells a poor cDNA starting source. In addition, the isolation of a murine cathepsin K genomic clone is described.
15   To isolate a genomic clone containing the cathepsin K gene, a human PAC library (Genome Systems, St. Louis, MO) was screened by PCR using a cathepsin K sequence tagged site (STS) from the 3' untranslated region (UTR). PAC DNA from the positive clone, 74e16, was isolated by standard methods.
20 Using the genomic organization data from cathepsins S and L (Shi, G.P. et al., 1994, J. Biol. Chem. 269:11530-11536), probable sites for exon-intron boundaries in the cathepsin K cDNA were selected. The seven predicted introns were PCR amplified from either human genomic or PAC 74e16 DNA with
25 exonic oligonucleotides. The amplified products were sized on an agarose gel and cycle sequenced on an ABI 377 sequencer. The genomic sequence was divided into eight exons ranging from 103 to 665 nt (FIG. 4). Exon 2 contained the translation initiation site, all codons for the signaling
30 prepeptide and a portion of the propeptide, exon 4 encoded the C-terminus of the propeptide and the N-terminus of the mature enzyme, and exon 8 contained the stop codon. The seven introns ranged in size from 86 nt to 2.4 kb, the entire gene being approximately 9 kb.
35   As shown in Table 1, below, the exon/intron boundary sequences conformed to the GT/AG rule with a single exception (Shapiro, M.B. and Senapathy, P., 1987, Nucl. Acids Res.

- 61 -

15:7155-7174). The 5' splicing donor site of exon 3 was non-conforming, containing the sequence TGgc, a pattern observed in approximately 1% of these sites. The boundary sequences were generally consistent with the 5' and 3' splice consensus

5 sequences. Variations from the consensus sequences were all permissible for the primate exon/intron boundary sequences.

10

15

20

25

30

35

TABLE 1

Exon/Intron Boundaries of the Human Cathepsin K Gene

Exon number and (size)	cDNA position of exon	5' Splicing site	Intron number and (size)	3' splicing site	Codon Phase
(nt)			(kb)		
1 (103)	-104--2	TCAGCAG gtaacg	1 (1.2)	ctgtttccctgccaaatgggagagtttccctaactacattcttctgcag GATGTGGGG	II
2 (122)	-1-121	CAACAAG gtgcct	2 (0.463)	catatgtaactgtagacagtctatacaagtactgactatgctttgtttag GTGGATGAA	I
3 (123)	122-244	GGACATG gcaagt	3 (0.086)	atttgcttagttccctgctgatgctggcctcttctttctttgtcttag ACCAGTGAA	I
4 (156)	245-400	AAATCAG gtactc	4 (1.5)	atagaaatgtaaacagcaaagattgatagtttctctgtatgcctttcag GGTCAGTGT tttttcttcncag G	I
5 (219)	401-619	GGGACAG gt gagaga (a)	5 (0.64)	tcattgctattgcttgtccctagtcctattacaatctgttcttccag GAAGAGAGT	I
6 (166)	620-785	AGCAAAG gtaaga	6 (0.270)	gggtgacccttggattgcatagagcctcacgctggtagtttgtattctag GTGTGTATT	III
7 (106)	786-891	AAAACAG gtaatg	7 (2.4)	tgaataaataaactgagagtactaagtatttttcttgattggtcttacag CTGGGGAGA	II
8 (665)	892-1556				

Consensus sequence: donor: C

acceptor: ccccccctccc

To determine the cathepsin K transcription initiation
site, primer extension analysis was performed using a
cathepsin K-specific antisense oligonucleotide primer which
was endlabeled with $[^{32}P]-\gamma$-ATP using T4 polynucleotide
5 kinase. Either human brain total RNA or human lung mRNA was
used with the radiolabeled primer for the reverse
transcription reaction. The primer extension experiments
were carried out as previously described (Shi et al., supra).
Three bands, corresponding to positions -169, -102, and -40,
10 were identified (FIG. 5).
    The 5' flanking region for the cathepsin K gene was
subcloned into pGEM4z from PAC 74e16 after Nsi I digestion
and identified by colony hybridization with an end-labeled
exon 1 oligonucleotide. Comparison of this 5' flanking
15 sequence cloned with the 5' UTRs from the cathepsin K cDNAs
revealed sequence divergence (FIG. 3). The 18 nt 5' UTR from
the osteoclastoma cathepsin k cDNA (Li, Y.-P. et al., 1995,
J. Bone Miner. Res. 10:1197-1202) matched all other
sequences, and is presumed to represent a partial transcript.
20 The 5' UTR of a macrophage-derived cDNA (Shi, G.P. et al.,
1995, FEBS Lett. 357:129-134) was 105 nt in length and
matched the genomic sequence, nearing the transcription start
site at -40. Its first 12 nt did not match and presumably
represents a cloning artefact. The cathepsin K 5' UTR cloned
25 from arthritic bone (Inaoka, T. et al., 1995, Biochem.
Biophys. Res. Commun. 206:89-96) showed identity with 46 nt
of exon 1 but its first 84 nt did not match the genomic
sequence. Sequence comparison of these 84 nt revealed
complete identity with the cathepsin K coding sequence at nt
30 864-876 and the reverse complement of 786-856.
Interestingly, this corresponded to the 5' boundary of exon
7, suggesting that this 5' UTR sequence was not an cloning
artefact but rather an error in transcription and/or message
splicing.
35   In contrast, the 141 nt 5' UTR from a spleen cathepsin K
cDNA clone reported by Bromme and Okamoto (Bromme, D. et al.,
1995, Biol. Chem. Hoppe-Seyler 376:379-384.) was identical to

100 nt of exon 1, ending at the -102 transcriptional start
site, but failed to match the genomic sequence at its most 5'
40 nt (FIG. 6). BLASTn searching of the non-redundant
sequence and dbEST databases with these 40 nt matched only
5   self. To determine the authentic 5' sequence, sense
oligonucleotides, corresponding to this 5' UTR and to the
predicted sequence near the transcription initiation site
(primer CTSK43 in FIG. G), and an antisense primer in exon 2
were used to PCR amplify the 5' UTR from human brain and lung
10 randomly primed first-strand cDNAs. A product of the
predicted size was amplified using the primer matching the
genomic sequence, but no product was amplified using the
primer based on the Bromme-Okamoto 5' UTR. These results
demonstrated that the 5' flanking region shown in FIG. 6 is
15 the major, and perhaps only, promoter in lung and brain.
Since the relevant portions of cathepsin K cDNA clones from
bone and osteoclastomas showed sequence identity with exon 1,
this 5' flanking region is likely the promoter controlling
expression of cathepsin K in osteoclasts as well.
20   Analysis of the genomic sequence upstream from the
transcription initiation site revealed a rather featureless
promoter which lacked the canonical TATA and CAAT box
sequences. Promoters from other members of the papain family
of cysteine proteinases, cathepsins S and B, have also lacked
25 TATA boxes (Gong, Q., et al. 1993, DNA Cell Biol. 12:299-309;
Shi, G.P. et al., 1994, J. Biol. Chem. 269:11530-11536). The
cathepsin K promoter contained two AP1 sites at -315 and -
346, was not particularly GC rich and lacked SP1 sites. This
promoter was similar to the cathepsin S promoter which
30 contained multiple AP1 sites and few SP1 sites but contrasted
with the cathepsin B promoter which was GC rich and contained
15 SP1 sites, corresponding well with the fact that both
cathepsins K and S are highly regulated while cathepsin B is
expressed constitutively. The signals which regulate
35 cathepsin K expression are not known. Since the principal
role for cathepsin K is in bone remodeling, putative estrogen
response elements were specifically sought in this genomic

sequence but none was found. Such elements have been
identified further upstream in promoters or in introns in
other estrogen-responsive genes (Sohrabji, F., et al., 1995,
Proc. Natl. Acad. Sci. U.S.A. 92:11110-11114), indicating a
5 possible role of estrogen in regulating the expression of
cathepsin K in osteoclasts.

The chromosomal localization of the cathepsin K gene was
identified by FISH. Purified human cathepsin K cDNA was
labeled with biotin 11-dUTP (Gibco BRL, Gaithersburg, MD) by
10 nick translation and hybridized to human metaphase
chromosomes from two normal males as previously described.
Probe signals were detected with avidin-fluorescein antibody
(X,X) and chromosomes were stained with 4,6-diamidino-2-
phenylindole dihydrochroride (DAPI; Oncor, Gaithersburg, MD).
15 Among the 22 metaphases scored, hybridization signal was
observed at chromosome band 1q21 from one or both chromatids
in 21. Since cathepsin S had also been mapped to 1q21 by
FISH (Shi, G.P. et al. 1994, J. Biol. Chem. 269:11530-11536),
the possibility was considered that cathepsins K and S might
20 reside in tandem. Indeed, STSs for both cathepsin K and S
were amplified from PAC 74e16 by PCR. Since PAC insert sizes
range from 100-150 kb, these two cathepsin genes, each
approximately 10 kb in length, lie in close proximity to one
another.

25   In summary, the cathepsin K gene was localized to within
150 kb of cathepsin S on chromosome 1q21. Like cathepsins S
and L with which it is highly homologous, the cathepsin K
genomic structure comprised 8 exons that were interrupted by
introns in the same coding sequence, suggesting that all
30 three genes evolved from a common ancestral cysteine
proteinase gene. Three transcription initiation sites were
identified at nt -164, -102, and -40 human brain and lung.
Analysis of the 5' flanking region, believed relevant for the
high expression in osteoclasts, revealed a TATA-less promoter
35 with two AP1 sites. The knowledge of the cathepsin K genomic
structure will, for example, facilitate mutation analysis in
pycnodysostosis patients and further studies of the

evolutionary relatedness of the lysosomal cysteine protease
genes.

9.   EXAMPLE: ISOLATION AND CHARACTERIZATION OF THE MURINE
     CATHEPSIN K GENE

In this Example, the isolation and characterization of
the murine cathepsin K gene is presented. The isolation of a
murine cathepsin K gene will facilitate the development of a
murine model of cathepsin K deficiency.

9.1 MATERIALS AND METHODS

9.1.1. Amplification of a mouse cathepsin K cDNA
fragment. Sense and antisense oligonucleotide primers based
on a highly conserved region of the human and rabbit
cathepsin K gene sequences C, were synthesized on an Applied
Biosystems (ABI) Model 394 DNA Synthesizer. These primers
were used to PCR-amplify a fragment from first strand cDNA
which was generated from mouse spleen total RNA by reverse
transcription. This 462-bp fragment was isolated and its
sequence proved to be highly similar to those of the known
cathepsin K sequences. This mouse partial cDNA was then used
to construct new sense (5'-TGTTGGGCTTTCAGCTCTGC-3') and
antisense (5'-CTGGGCCACATATGGGTAAGCAT-3') primers expected to
reside within exon 5 based on the human cathepsin K genomic
organization. This second set of primers amplified a 205-bp
fragment from mouse genomic DNA.

9.1.2 Isolation and sequencing of mouse-cathepsin K
cDNA clones. Since Northern analysis revealed moderate
cathepsin K transcript levels in human heart (Brömme, D. et
al., 1995 Biol. Chem. Hoppe-Seyler 376:379-384), a BALB/C
mouse heart cDNA phagemid library (Stratagene, La Jolla, CA)
was screened with a mouse cathepsin K probe. The
radiolabeled probe was produced by encorporating [α-^{32}P]dCTP
(Amersham, Arlington Heights, [L) during PCR with the exon 5
primers. Positive plaques were selected and unique plaques
were isolated by secondary screening. Insert-containing

plasmid was excised from the phagemid using M13 helper phage
according to the manufacturer's instructions (Stratagene).
After a standard plasmid DNA minipreparation, the inserts
were released from the pBluescriptSK vector by restriction
5 digestion with EcoRI (New England Biolabs, Beverly, MA) and
then sized by electrophoresis in a 1% agarose gel. The clone
containing the largest insert was selected for sequence
analysis. Initial DNA sequencing was performed with T7 and
T3 primers using an ABI 377 Sequencer. Subsequent reverse
10 orientation sequencing was performed using insert-specific
primers. Analyses of the sequence data were performed with
the AutoAssembler (ABI, Perkin Elmer, Foster City, CA) and
MacVector (Kodak, Rochester, NY) software packages.

15   9.1.3 Amplification and sequencing of the 5' region of
the mouse cathepsin K cDNA. In order to obtain sequence for
the 5' portion of the mouse cathepsin K coding region, 5'
rapid amplification of CDNA ends (RACE) (Frothman, M.A., et
al., 1988, Proc. Natl. Acad. Sci USA 85:8998-9002) was
20 performed according to the manufacturer's instructions
(Bochringer Mannheim, Indianapolis, IN). Total RNA extracted
from mouse melanocyte cultures was reverse-transcribed using
MMLV reverse transcriptase and the mouse cathepsin K-specific
antisense oligonucleotide from exon 5. The first-strand cDNA
25 products were tailed using terminal transferase
deoxyadenosine. Two rounds of PCR were used to amplify the
5' region using nested mouse cathepsin K-specific
oligonucleotide primers (first round: 5'-
CAAGTGGTTCATGGCCAGTTC-3'; second round: 5'-
30 CCATAGCTCCCACTGGGTGT-3!). The amplified products were sized
on an agarose gel by horizontal electrophoresis, extracted
from agarose slices, and then sequenced by cycle sequencing
on the ABI 377 Sequencer.

35

9.1.4   Characterization of the mouse cathepsin K genomic organization.   Using the genomic organization data from human cathepsins K, S, and L (Shi, G.P. et al., 1994 J. Biol. Chem. 269:11530-11536) probable sites for exon-intron boundaries in
5   the mouse cathepsin K cDNA were identified.   Sense and antisense exonic primers flanking introns 2 through 7 were synthesized.   Mouse genomic clones containing the cathepsin K gene were obtained by screening a SVI29 mouse X bacteriophage library (Stratagene) using mouse cathepsin cDNA that had been
10   radiolabeled with [α-^{32}P]dCTP (Amersham) by the random hexamer method according to the manufacturer's instructions (New England Biolabs).   Individual positive clones were isolated and phage DNA was isolated by standard methods.   Southern analysis of the clones was performed, using exonic
15   oligonucleotides that had been end-labelled with [γ-32P]ATP (Amersham) by polynucleotide kinase (New England Biolabs).   Introns 2-7 were PCR-amplified from the phage DNA with exonic oligonucleotides and the amplified products were sized on an agarose gel.
20

9.2   RESULTS
9.2.1   Isolation and Characterization of the mouse Cathepsin K cDNA.   The original cDNA fragment amplified from mouse spleen RNA was 462 bp in length and was highly
25   homologous to the other known cathepsin K sequences (e.g., this fragment had 86% identity to the human sequence).   Screening of the mouse heart cDNA library yielded eight partial cathepsin K cDNA clones.   The longest clone, pMCatKI, contained the complete coding sequence for the mature
30   peptide, the entire 3' untranslated region (UTR), and the sequence coding for 97 of the 99 amino acid residues expected for the prepeptide.   pMCatK1 lacked the 5' UTR, the initiation ATG, the coding region for the signaling prepeptide and sequence for the first two residues of the
35   prepeptide.
        By 5' RACE, a 250-bp fragment was amplified from the mouse melanocyte RNA and sequenced.   This fragment contained

the remainder of the coding region and 135 bp of the 5' UTR.
Comparison of this murine sequence with the human cathepsin K
cDNA revealed some homology in the 5' UTR and near identity
in the coding region. BLAST searching of the dbEST database
5 detected four murine ESTs (AA107496, W13542, W14460, and
AA097791). These ESTs were identical in sequence to both the
RACE fragment and pMCatK1 across the portions of the 5' UTR
.and coding region which they contained, confirming that the
RACE fragment derived from the mouse cathepsin K gene.
10   The sequence for the partial 5' UTR, the entire coding
sequence, and tile 3' UTR with the 9-bp poly(A) tract along
with the predicted protein sequence is shown in Figure 7.
The initiation ATG was preceded by a Kozak consensus site
(Kozuk, M. et al., 1986, Cell 44:283-292). No other ATG
15 codons were present in the 5' UTR sequence. The coding
region was 990 bp in length and encoded a 15-residue
prepeptide, a 99-residue pro-piece, and a 215-amino-acid
mature enzyme. Two potential glycosylation sites, one in the
pro-piece (at codons N103, D104, T105) and another in the
20 mature enzyme (at codons N213, A214, T215) were detected.
The 3' UTR was 46I bp in length, not including the poly(A)
tract. There were two putative consensus polyadenylation
signals (AATAAA) located 42 and 18 bp upstream from the
poly(A) tract. In addition, a single consensus recognition
25 sequence (CACTG) for the U4 small nuclear ribonucleoprotein
(Berget, S.M., 1984, Nature 309:179-182) was located 126 bp
upstream of the poly(A) tract.
        Comparison of the mouse cathepsin K cDNA with the human
and rabbit homologues revealed 87% nucleotide identity across
30 the entire coding region in both the human and rabbit
sequences with similar conservation in the pre-, pro-, and
mature peptide regions. Comparison of the mouse and chicken
cathepsin K cDNAs demonstrated an overall 70% nucleotide
identity but only 40% nucleotide identity across the
35 prepeptide region which also had a 15-nucleotide insert
encoding five additional amino acids in the chicken gene.
Amino acid sequence alignment of the mouse, human, rabbit,

and chicken cathepsin K prepolypeptides (FIG. 8) revealed 85%
identity and 93% similarity between the mouse and human amino
acid sequences, 88% identity and 94% similarity between mouse
and rabbit, and 72% identity and 86% similarity between the
5 mouse and chicken prepropeptides.

9.2.3   Partial Characterization of the Mouse
             Cathepsin K Genomic Structure

Six λ bacteriophage clones were identified and isolated
10 from the SV129 genomic library. Southern analysis with end-
labeled oligonucleotides corresponding to sequences from
exons 2 and 8 detected bands summing to approximately 9 kb.
Exons 2-6 were shown to reside on a single 6-kb HindIII
fragment, while exons 6-8 were on a 3 kb HindIII fragment.
15 These data suggested that, like the human cathepsin K genomic
structure which was shown to be approximately 9 kb in its
complete length, the murine genomic homologue was similar in
size.

Using murine exonic PCR primers based on the location of
20 the exons in the human cathepsins K, L, and S genes, the
intervening introns in the mouse cathepsin K gene were
successfully amplified and sized. The putative structure of
the mouse cathepsin K gene is shown in FIG. 9. Comparison of
the mouse and human cathepsin K genomic structures revealed
25 substantial conservation of intron sizes. In addition,
limited sequencing, of intron 2 documented the presence of a
dinucleotide repeat, as had previously been found at the same
position in the human gene.

The present invention is not to be limited in scope by
30 the exemplified embodiments which are intended as
illustrations of single aspects of the invention, and any
clones, DNA or amino acid sequences which are functionally
equivalent are within the scope of the invention. Indeed,
various modifications of the invention in addition to those
35 described herein will become apparent to those skilled in the
art from the foregoing description and accompanying drawings.
Such modifications are intended to fall within the scope of

the appended claims. It is also to be understood that all
base pair sizes given for nucleotides are approximate and are
used for purposes of description.

5

10

15

20

25

30

35

WHAT IS CLAIMED IS:

1.   A method for ameliorating bone resorption disorder
symptoms, comprising: contacting a compound capable of
5 specifically inhibiting cathepsin S activity to an osteoclast
for a time sufficient to inhibit the cathepsin K activity of
the osteoclast so that symptoms of the bone resorption
disorder are ameliorated.

10   2.   The method of Claim 1, further comprising:
contacting a second compound to the osteoclast, wherein the
second compound is capable of specifically inhibiting a
second bone resorption activity for a time sufficient to
inhibit the cathepsin K and the second activity of the
15 osteoclast so that symptoms of the bone resorption disorder
are ameliorated.

3.   The method of Claim 1, wherein the compound further
inhibits a second bone resorption activity.
20

4.   The method of Claim 2 wherein the second activity
is cathepsin S activity.

5.   The method of Claim 3 wherein the second activity
25 is cathepsin S activity.

6.   A method for ameliorating macrophage-mediated
inflammatory damage symptoms, comprising: contacting a
compound capable of specifically inhibiting cathepsin S
30 activity to an macrophage for a time sufficient to inhibit
the cathepsin K activity of the macrophage so that symptoms
of the macrophage-mediated inflammatory damage are
ameliorated.

35   7.   The method of Claim 6, further comprising:
contacting a second compound to the macrophage, wherein the
second compound is capable of specifically inhibiting a

second macrophage-mediated inflammatory damage activity for a
time sufficient to inhibit the cathepsin K and the second
activity of the macrophage so that symptoms of the
macrophage-mediated inflammatory damage are ameliorated.

5

8.   The method of Claim 6, wherein the compound further
inhibits a second macrophage-mediated inflammatory damage.

9.   The method of Claim 7 wherein the second activity
10 is cathepsin S activity.

10.   The method of Claim 8 wherein the second activity
is cathepsin S activity.

15   11.   A transgenic animal which carries at least one

cathepsin transgene whereby said cathepsin transgene is

expressed or overexpressed.

20

      12.   The transgenic animal of Claim 11 wherein the

cathepsin transgene is a cathepsin genomic or cDNA transgene.

      13.   The transgenic animal of Claim 12 wherein the
25
cathepsin genomic or cDNA transgene is murine or human.

      14.   The transgenic animal of Claim 11 wherein the
30 animal is a mouse.

      15.   The transgenic animal of Claim 11 wherein the

cathepsin gene encodes one of the following: cathepsin K or

35 cathepsin S.

16.　The transgenic animal of Claim 11 which carries a cathepsin K and cathepsin S transgene.

5　　　17.　A transgenic animal which is deficient for at least one cathepsin gene.

18.　The transgenic animal of Claim 17 wherein the
10 animal is a mouse.

19.　The transgenic animal of Claim 17 wherein the deficient cathepsin transgene is cathepsin K or cathepsin S.
15

20.　The transgenic animal of Claim 17 wherein the deficient cathepsin transgene is cathepsin K and cathepsin S.

20

25

30
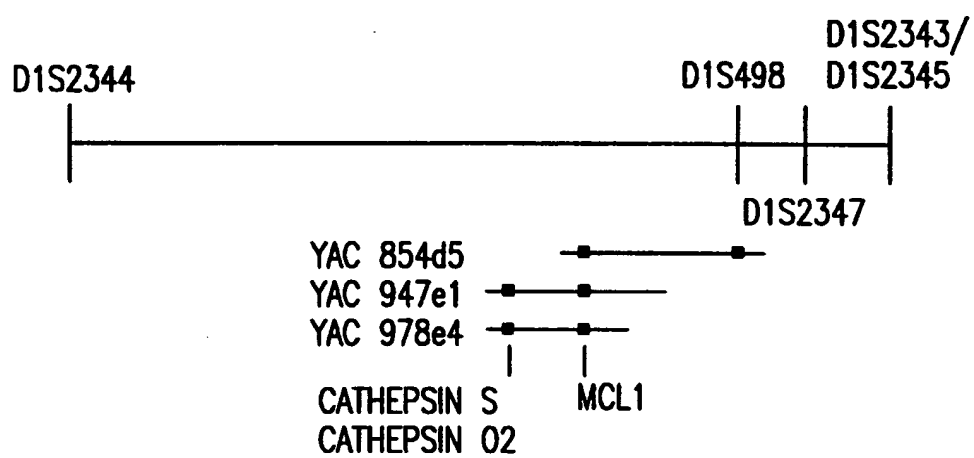

35

FIG.1
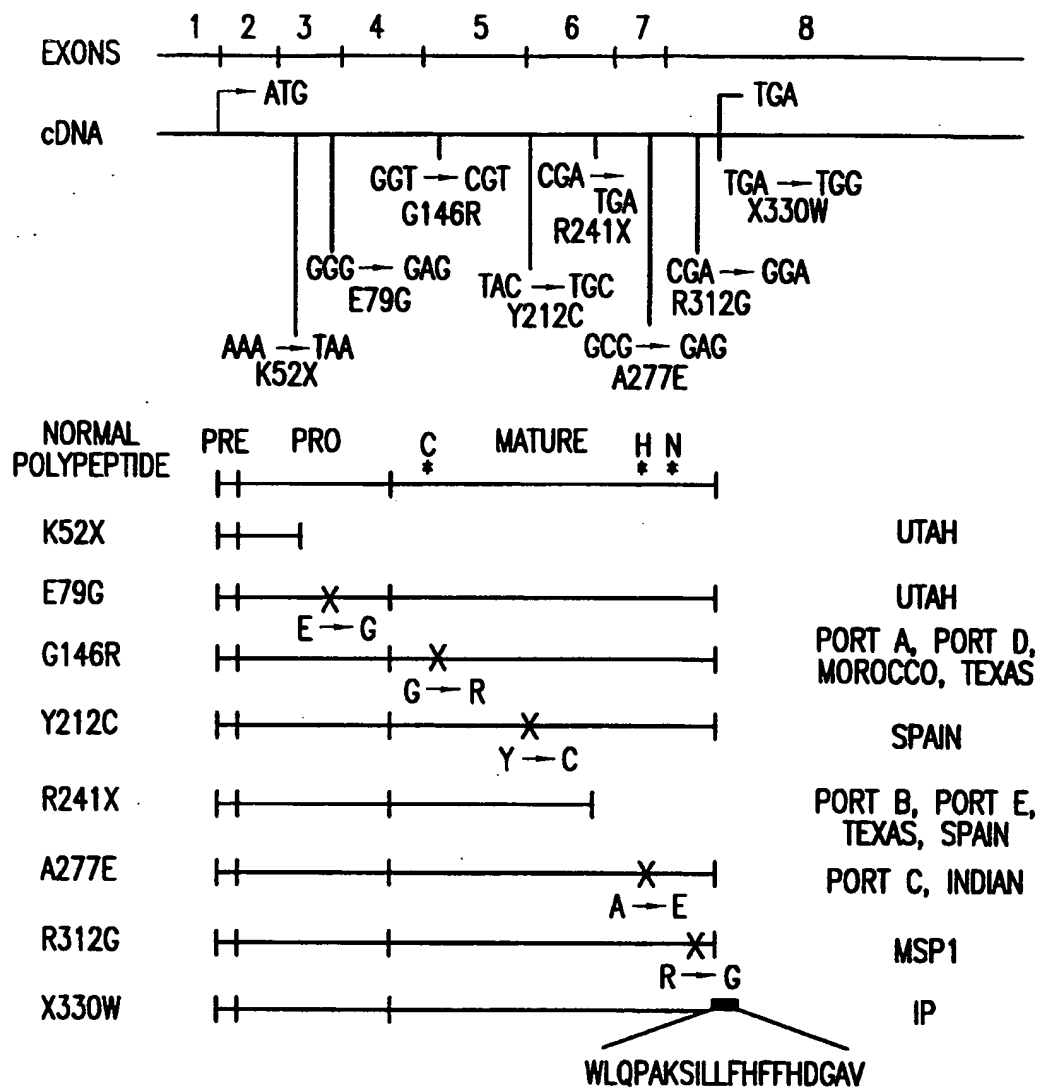
FIG.2

cDNA: --- W M W M W M

Cathepsin K — — — — — 28 K

— 18 K

1 2 3 4 5 6 7

FIG.3A

cDNA: -- W M W M W M

←Cathepsin K

←VAI

1 2 3 4 5 6 7

FIG.3B

FIG.4

PE G A T C

253 bp →

186 bp →

124 bp →

FIG.5

```
       -592  ccaacaagcttataaaaatacaaagaattctggagcctatggcttccattacattactttttatagcctttatgttca  -513
       -512  ttaccgcatccagaggtgagagtcagacacaaatgaaaatagttcaatgttggagaggtaaatcctaacaggaaa  -433
       -432  ggggtaggaaaagatataatccccaatattaaaataaagatattgaagaagaaggatgggagagactaggctgtgtcc  -353
       -352  ttccttttactcaccaaagagaaagtaagctcctattggagtcaatagatattgggtcttgttattgccaccaaga  -273
                                                                    AP1
       -272  cagtcttgtgagactaaatagctagtaatccctaccctggcacacatgctgcatacacacagaaacactgcaaatccac  -193
                AP1
Genomic -192  tgcctcctтcctcctccctacCCTTCTCTCTCAGCATTCTATCCCCGCTCCTCCTCTCTTACCCAAATTTCCAGCC  -113
                                                                              CTSK43
                                                                          ──────→
Spleen cDNA:                    gcgcactcacagtcgcaaccttttcccttc
                                                                            A
Bone cDNA:                      gaaacaggcactggattccatatcccctgcccaaaa
                                                          ─────────
                                                              B

Genomic -112  GATCACTGGAGCTGACTTCCGCAATCCCGATGGAATAAATTCTAGCACCCCTGATGGTGTGCCCACTTGCTGCCGAA  -33
                           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Spleen cDNA:               ctgacttccgcttccgcaatccgaatccgatggaataaattctagcaccccctgatggtgtgcccacactttgctgccgaa
                                    |||||||||||||||||||||||||||||||||||||||||||||||
Macrophage cDNA:          attcggcacgagccgcaatccgatggaataaattctagcaccccctgatggtgtgcccacactttgctgccgaa
                                    |||||||||||||||||||||||||||||||||||||||||||||||
Bone cDNA:                ccgcatggttcagattatcgctattgcagctttcatcatcataatacacacctttgctcatcataatacacacctttgctgccgaa
                                                                                      B

Genomic  -32  ACGAAGCCAGACAACAGATTTCCATCAGCAG  -2
                           |||||||||||||||||||||||||||||||
Spleen cDNA:              acgaagccagacaacagatttccatcagcag
                           |||||||||||||||||||||||||||||||
Macrophage cDNA:          acgaagccagacaacagatttccatcagcag
                           |||||||||||||||||||||||||||||||
Bone cDNA:                acgaagccagacaacagatttccatcagcag
                                         ||||||||||||||||
Osteoclastoma cDNA:                     cagatttccatcagcag
```

FIG.6

TCCAGATTTTCCAGC CACTGCTGGAGTTGACTTCC GCAATCCTTACCGAATAAAT   55

TTAGCACCCTTAGTCTTCCG CTCACAGTAGCCACGCTTCT TATCCGAAAAGAGCCTAGCG AACAGATTCTCAACAGCAGG   135

Pre   Pro
ATG TGG GTG TTC AAG TTT CTG CTG CTA CCC ATG GTG AGC TTT GCT CTG TCT CCG GAG GAA ATG   198
Met Trp Val Phe Lys Phe Leu Leu Leu Pro Met Val Ser Phe Ala Leu Ser Pro Glu Glu Met

1

CTG GAC ACC CAG TGG CTA GAG CTA TGG AAG ACT CAC CAG AAG CAG TAT AAC AGC AAG GTG GAT   261
Leu Asp Thr Gln Trp Leu Glu Leu Trp Lys Thr His Gln Lys Gln Tyr Asn Ser Lys Val Asp

22

GAA ATC TCT CGG CGT TTA ATT TGG GAG AAA AAC CTG GAG AAG CAA ATC TCT GCC CAT AAC CTG GAG   324
Glu Ile Ser Arg Arg Leu Ile Trp Glu Lys Asn Leu Glu Lys Gln Ile Ser Ala His Asn Leu Glu

43

GCC TCT CTT GGT GTC CAT ACA TAT GAA CTG GCC ATG AAC CAC TTG GGA GAC ATG ACC AGT GAA   387
Ala Ser Leu Gly Val His Thr Tyr Glu Leu Ala Met Asn His Leu Gly Asp Met Thr Ser Glu

64

GAA GTG GTT CAG AAG ATG ACG GGA ATA CCT CCC TCT CGA TCC TAC AGT TAT AAT GAC ACT   450
Glu Val Val Gln Lys Met Thr Gly Ile Pro Pro Ser Arg Ser Tyr Ser Tyr Asn Asp Thr

85

Mature
CTC TAT ACC CCA GAG TGG GAA GGC AGG GTC CCA GAC TCC ATC GAC TAT CGA AAG AAA GGA TAC   513
Leu Tyr Thr Pro Glu Trp Glu Gly Arg Val Pro Asp Ser Ile Asp Tyr Arg Lys Lys Gly Tyr

106

GTT ACT CCA GTC AAG AAC CAG GGC CAG TGT GGT TCC GGT CAG GGG GCC TCT GCC TTC AGC GCC GGG GCC   576
Val Thr Pro Val Lys Asn Gln Gly Gln Cys Gly Ser Cys Trp Ala Phe Ser Ala Gly Ala

127

CTG GAG GGC CAA CTC AAG AAA ACT GGT CTT AAA CTC TTA GCT CTG AGT CCC CAG AAT CTT GTG   639
Leu Glu Gly Gln Leu Lys Lys Thr Gly Leu Lys Leu Leu Ala Leu Ser Pro Gln Asn Leu Val

148

FIG.7A

```
169  GAC TGT GTG ACT GAG AAT TAT GGC TGT GGA GGC GGC TAT ATG ACC ACT GCC TTC CAA TAC GTG      702
     Asp Cys Val Thr Glu Asn Tyr Gly Cys Gly Gly Gly Tyr Met Thr Thr Ala Phe Gln Tyr Val

190  CAG CAG AAC GGA GGC ATT GAC TCT GAA GAT GCT CCA TAC GTG GGC CAG GAT GAA AGT TGT           765
     Gln Gln Asn Gly Gly Ile Asp Ser Glu Asp Ala Tyr Pro Val Gly Gln Asp Glu Ser Cys

211  ATG TAT AAC GCC ACG GCA AAG GCA GCT AAA TGC AGA GGG TAC GAG ATT CCT GTG GGG AAC           828
     Met Tyr Asn Ala Thr Ala Lys Ala Ala Lys Cys Arg Gly Tyr Arg Glu Ile Pro Val Gly Asn

232  GAC AAA GCC CTG AAG AGA GCA GTG GCG CGG GTA GGA CCC ATC TCT GTG TCC ATC GAT GCA AGC       891
     Glu Lys Ala Leu Lys Arg Ala Val Ala Arg Val Gly Pro Ile Ser Val Ser Ile Asp Ala Ser

253  TTG GCA TCT TTC CAG TTT TAC AGC AGA AGT GTG TAC TAT GAT GAA AAT TGT GAC CGT GAT AAT       954
     Leu Ala Ser Phe Gln Phe Tyr Ser Arg Ser Val Tyr Tyr Asp Glu Asn Cys Asp Arg Asp Asn

274  GTG AAC CAT GCA GTG TTG GTG GTG GGC TAT GGC ACC CAG AAG GGA AGC AAG CAC TGG ATA ATT      1017
     Val Asn His Ala Val Leu Val Val Gly Tyr Gly Thr Gln Lys Gly Ser Lys His Trp Ile Ile

295  AAA AAC AGC GGA GAG AGC TGG GGA AGC TAT CTC TTG GCT CTC TTG GCT CGG AAT AAG AAC          1080
     Lys Asn Ser Gly Glu Ser Trp Gly Ser Tyr Leu Ala Leu Ala Arg Asn Lys Asn

316  AAC GCC TGC GGC ATT ACC PAC ATG CCC AAG TTC CCC AGC TTC ATG TGA TTCCAGCCAGCCAGCCATCTCT  1148
     Asn Ala Cys Gly Ile Thr Asn Met Pro Lys Phe Pro Ser Phe Met End
```

```
TCTCAGATTCCTTCCTTCAT CGTGCAAGATATTGGTGGCT TTCGAAGGGAGTGGGCATGG CGGCTCCTGAGAGAGACGCAG   1228
CGATGCTAACTAAGATTGTT TGATTCTCCTCGTTGGT GCTTCCAGTGACAACTCTAC TTCCTTCTCTCTGCCCAGGC      1308
CCCTTTCTTTGTTGGACACA ACAGGGATTTTTCACAGTT GTGGACTCTCTGCTCGTAGA CATTGGACTCCTCCACCAGG      1388
CTGGAGGACTAAGGTGACCT TCCCGAGCCCCTCTCTTCTC TATACACCAGTGAACATTTC AGTCTTCCACTGACATCCAC      1468
AAATCTATTCGTGATTCTTT GACAAATTTACATCATATTA AAAAAGGTGTTTTTCCTTCT TTGTATTTGAAATAAAGTAT      1548
CTCATTTACAATTTAATAAA GAACACTGCCTCCATACTAA AAAAAAA                                       1595
```

FIG.7B

```
                                                          45
Mouse    M-----WVFK FLLLPMVSFA LSPEEMLDTQ WELWKKTHQK QYNSKVDEIS
Human    M-----WGLK VLLLPVVSFA LYPEEILDTH WELWKKTHRK QYNNKVDEIS
Rabbit   M-----WGLK VLLLPVVSFA LHPEEILDTQ WELWKKTYSK QYNSKVDEIS
Chicken  MLRLHWLALL VLLLPMAAAQ LRPEPELDAQ WDLWKRTIQK QYNGKEDETS

                                                          95
Mouse    RRLIWEKNLK QISAHNLEAS LGVHTYELAM NHLGDMTSEE VVQKMTGLRI
Human    RRLIWEKNLK YISIHNLEAS LGVHTYELAM NHLGDMTSEE VVQKMTGLKV
Rabbit   RRLIWEKNLK HISIHNLEAS LGVHTYELAM NHLGDMTSEE VVQKMTGLKV
Chicken  RRLIWEKNLK YIDAHNAEAR LGKHSFQLAM NYLGDMTSEE VVRTMTGLRV

                                                  *       145
Mouse    PPSRSYSNDT LYTPEWEGRV PDSIDYRKKG YVTPVKNQGQ CGSCWAFSSA
Human    PLSHSRSNDT LYIPEWEGRA PDSVDYRKKG YVTPVKNQGQ CGSCWAFSSV
Rabbit   PPSRSHSNDT LYIPDWEGRT PDSIDYRKKG YVTPVKNQGQ CGSCWAFSSV
Chicken  PRSRPRPNGT LYVPDWSSRA PAAVDWRRKG YVTPVKDQGQ CGSCWAFSSV

                                                          195
Mouse    GALEGQLKKK TGKLLALSPQ NLVDCVTENY GGGGGYMTTA FQYVQQNGGI
Human    GALEGQLKKK TGKLLNLSPQ NLVDCVSEND GGGGGYMTNA FQYVQKNRGI
Rabbit   GALEGQLKKK TGKLLNLSPQ NLVDCVSENY GGGGGYMTNA FQYVQRNRGI
Chicken  GALEGQLKRR TGKLLSLSPQ NLVYCVSNNN GGGGGYMTNA FEYVRLNRGI

                                                          245
Mouse    DSEDAYPYVG QDESCMYNAT AKAAKCRGYR EIPVGNEKAL KRAVARVXPI
Human    DSEDAYPYVG QEESCMYNPT GKAAKCRGYR EIPEGNEKAL KRAVARVGPV
Rabbit   DSEDAYPYVG QDESCMYNPT GKAAKCRGYR EIPEGNEKAL KRAVARVGPV
Chicken  DSEDAYPYIG QDESCMYSPT GKAAKCRGYR EIPEDNEKAL KRAVARIGPV

                                     *                    295
Mouse    SVSIDASLAS FQFYSRSVYY DENCDRDNVN HAVLVVGYGT QKGSKHWIIK
Human    SVAIDASLTS FQFYSKGVYY DESCNSDNLN HAVLAVGYGI QKGNKHWIIK
Rabbit   SVAIDASLTS FQFYSKGVYY DENCSSDNVN HAVLAVGYGI QKGNKHWIIK
Chicken  SVGIDASLPS FQFYSRGVYY DTGCNPENIN HAVLAVGYGA QKGTKHWIIK

         *                           319
Mouse    NSWGESWGNK GYALLARNKN NACGITNMAS FPKM
Human    NSWGENWGNK GYILMARNKN NACGIANLAS FPKM
Rabbit   NSWGESWGNK GYILMARNKN NACGIANLAS FPKM
Chicken  NSWGTEWGNK GYVLLARNMK QTCGIANLAS FPKM
```

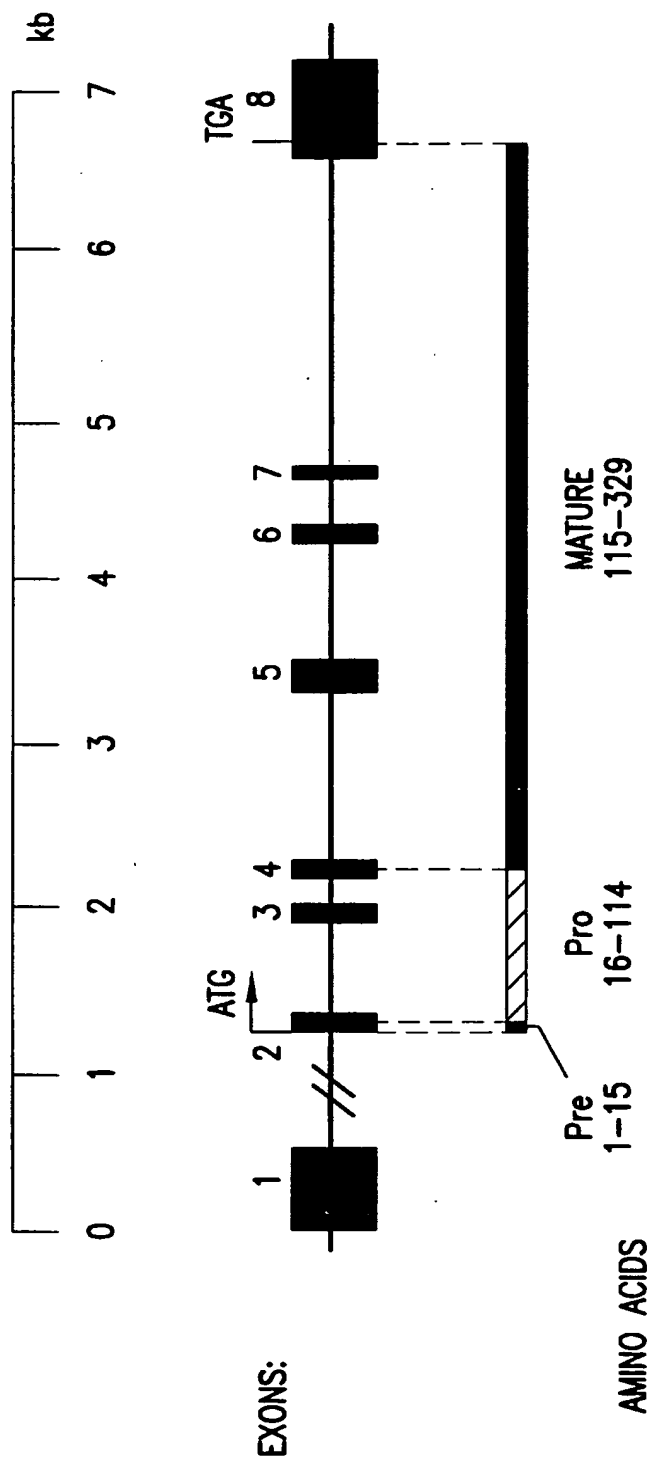# FIG.8

SUBSTITUTE SHEET (RULE 26)

FIG.9

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :A61K 31/00, 38/00, 48/00; C12N 5/00, 15/00

US CL :514/1, 2, 44; 800/2

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 514/1, 2, 44; 800/2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, CHEMICAL ABSTRACTS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	BOSSARD, M.J. et al. Proteolytic Activity of Human Osteoclast Cathepsin K. Journal of Biological Chemistry. 1996, Vol. 271, pages 12517-12524, especially pages 12520-12523.	1-20NON
Y	INAOKA, T. et al. Molecular Cloning of Human cDNA for Cathepsin K: Novel Cysteine Proteinase Predominantly Expressed in Bone. Biochemical Biophysical Research Communication. 1995. Vol. 206, pages 89-96, especially pages 90-94.	1-20
Y	MUNGER, J. et al. Lysosomal Processing of Amyloid Precursor Protein to AβPeptides: A Distinct Role for Cathepsin S. Biochemical Journal. 1995. Vol. 311, pages 299-305, especially 303-304.	1-20

| X | Further documents are listed in the continuation of Box C. | ☐ | See patent family annex. |

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance		
"B"	earlier document published on or after the international filing date	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed	"&"	document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
07 JANUARY 1998	0 2 FEB 1998

Name and mailing address of the ISA/US	Authorized Officer
Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	DEBORAH CROUCH, PH.D.
Facsimile No. (703) 305-3230	Telephone No. (703) 308-0196

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	SHI, G. et al.Molecular Cloning and Expression of Human Alveolar Macrophage Cathepsin S, an Elastinolytic Cysteine Protease. Journal of Biological Chemistry. Vol. 267, pages 7258-7262, especially 7260-7261.	1-20
Y	CAPPECCHI, M. The New Mouse Genetics: Altering the Genome by Gene Targeting. Trends in Genetics. 1989. Vol. 5, pages 70-76, especially pages 71-73.	17-20
Y	GRISOLANO, J.L. et al. Early Myeloid Cell-Specific Exrpressing of the Human Cathepsin G gene in Transgenic Mice. Procedings of the National Academy of Sciences (USA). October, 1994. Vol. 91, No. 19, pages 8989-8993, especially pages 8991-8992.	11-16
Y	US 5,501,969 A (HASTINGS et al) 26 March 1996, col. 1, line 62 to col. 2, line 3.	1-10